

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2005, 08:29:23 ; Search time 1832 Seconds
(without alignments)
1243.120 Million cell updates/sec

Title: US-10-349-143-1

Perfect score: 47

Sequence: 1 ccacagctgattagaat.....gatcacatttcaatttt 47

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_srb:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46.6	99.1	47	6	AR288266 Sequence
2	46.6	99.1	93153	2	AF257499 Homo sapi
3	46.6	99.1	182151	9	AC011835 Homo sapi
4	46.6	99.1	185463	9	AC016065 Homo sapi
5	46.6	99.1	191377	9	AC018398 Homo sapi
6	46.6	99.1	240825	6	AX087869 Sequence
7	46.6	99.1	240825	6	AX523960 Sequence
8	29.4	62.6	142552	9	HS5188 Human DNA
9	29.4	62.6	160530	2	AC025631 Homo sapi
10	29.4	62.6	266079	2	AL391060 Homo sapi
11	27.6	58.7	40479	9	AL356492 Human DNA
12	27.6	58.7	70766	2	AC090635 Homo sapi
13	27.6	58.7	118335	8	AC006193 Arabidops
14	27.2	57.9	177483	9	AC108046 Homo sapi
15	27.2	57.9	225586	5	AX469934 Zebrafish
16	27.2	57.9	237042	2	AX908795 Danio rer
17	26.4	56.2	160796	10	AL592222 Mouse DNA
18	26.4	56.2	169994	5	BX31875 Zebrafish
19	26.4	56.2	197773	2	BX36440 Danio rer

20	26.4	56.2	203451	10	AC023173 Mus muscu
21	26.4	56.2	218449	10	AC122835 Mus muscu
22	26.2	55.7	213417	2	AC119515 Rattus no
23	26.2	55.7	217538	2	AC098034 Rattus no
24	26.2	55.7	242655	2	AC107465 Rattus no
25	26.2	55.7	313096	2	AC131879 Rattus no
26	26	55.3	138307	10	AL808014 Mouse DNA
27	26	55.3	162058	2	AC135447 Rattus no
28	26	55.3	238527	10	AC129700 Rattus no
29	26	55.3	244242	2	AC107551 Rattus no
30	26	55.3	321552	2	AC109528 Rattus no
31	25.8	54.9	128301	9	AC133480 Homo sapi
32	25.8	54.9	154109	2	CR376803 Danio rer
33	25.8	54.9	155754	2	AC011273 Homo sapi
34	25.8	54.9	162087	2	AC011576 Homo sapi
35	25.8	54.9	164423	2	CR361552 Danio rer
36	25.8	54.9	177203	2	AC024531 Homo sapi
37	25.8	54.9	204821	2	BX276111 Danio rer
38	25.8	54.9	225216	10	AL844888 Mouse DNA
39	25.6	54.5	1913	9	BC033135 Homo sapi
40	25.6	54.5	106302	9	AC080011 Homo sapi
41	25.6	54.5	150939	2	AC091676 Homo sapi
42	25.6	54.5	151386	2	AC015731 Homo sapi
43	25.6	54.5	160119	2	AC130400 Homo sapi
44	25.6	54.5	165407	2	AC025136 Homo sapi
45	25.4	54.0	6890	8	AB010274 Schizosac
46	25.4	54.0	23193	8	SPAC926
47	25.4	54.0	165825	2	BX511301 Danio rer
48	25.4	54.0	186858	10	AC130210 Mus muscu
49	25.2	53.6	65294	9	AC110791 Homo sapi
50	25.2	53.6	92587	8	AC096665 Homo sapi
51	25.2	53.6	93375	8	AC139853 Medicago
52	25.2	53.6	93375	8	AC139853 Medicago
53	25.2	53.6	95155	9	AC078950 Homo sapi
54	25.2	53.6	112823	2	AC013301 Homo sapi
55	25.2	53.6	150194	5	BX649510 Zebrafish
56	25.2	53.6	165349	2	AC093665 Homo sapi
57	25.2	53.6	166458	10	AC136455 Mus muscu
58	25.2	53.6	166857	2	AC098589 Homo sapi
59	25.2	53.6	199523	9	AC012014 Homo sapi
60	25.2	53.6	220892	2	AC111605 Rattus no
61	25	53.2	24521	10	BX649344 Mouse DNA
62	25	53.2	42271	2	AC100593 Mus muscu
63	25	53.2	42271	2	AC100593 Mus muscu
64	25	53.2	88982	5	AL732598 Zebrafish
65	25	53.2	110000	1	BX897699_06 Continuation (7 of
66	25	53.2	110000	1	BX897699_07 Continuation (8 of
67	25	53.2	110000	2	BX546462_1 Continuation (2 of
68	25	53.2	118016	10	AL929405 Mouse DNA
69	25	53.2	121466	9	AC105310 Homo sapi
70	25	53.2	166409	2	CR774190 Danio rer
71	25	53.2	177949	2	CR513782 Danio rer
72	25	53.2	179475	5	BX001025 Zebrafish
73	25	53.2	180320	10	AC127588 Mus muscu
74	25	53.2	181069	2	CR762471 Danio rer
75	25	53.2	183230	5	BX649488 Zebrafish
76	25	53.2	184923	2	CR847497 Danio rer
77	25	53.2	186379	2	CR405691 Danio rer
78	25	53.2	186645	10	AL807773 Mouse DNA
79	25	53.2	189787	2	CR293534 Danio rer
80	25	53.2	240548	2	AC094036 Rattus no
81	25	53.2	254657	5	BX004878 Zebrafish
82	25	53.2	257288	2	AC120720 Rattus no
83	25	53.2	328252	2	AC096800 Rattus no
84	24.8	52.8	47557	2	CR354592 Danio rer
85	24.8	52.8	66510	2	AC074386 Homo sapi
86	24.8	52.8	68130	3	AC024757 Homo sapi
87	24.8	52.8	85347	2	AC017560 Drosophila
88	24.8	52.8	89261	9	AC098863 Homo sapi
89	24.8	52.8	91268	8	AP004895 Lotus cor
90	24.8	52.8	99877	1	CP000003_18 Continuation (19 of
91	24.8	52.8	110000	2	AC151905 Mus muscu
92	24.8	52.8	122013	3	AC092339 Drosophila

```

c 93 24.8 52.8 128621 8 AC147002 Medicago
94 24.8 52.8 130030 9 AC004889 Homo sapi
95 24.8 52.8 132724 9 AC068713 Homo sapi
c 96 24.8 52.8 146059 5 AL929549 Zebrafish
97 24.8 52.8 151770 9 AC002066 Homo sapi
98 24.8 52.8 152467 2 AC140023 Medicago
99 24.8 52.8 152481 2 AC151210 Bos tauri
c 100 24.8 52.8 154904 9 AL162272 Human DNA

```

ALIGNMENTS

```

RESULT 1
AR288266 LOCUS AR288266 47 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 1 from patent US 6537751.
ACCESSION AR288266 GI:31675550
VERSION AR288266.1 GI:31675550
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..47
/organism="unknown"
/mol_type="genomic DNA"

```

```

ORIGIN
Query Match 99.1%; Score 46.6; DB 6; Length 47;
Best Local Similarity 100.0%; Pred. No. 0.0002;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 CCACAGGCTTGATTGAAATAAATTTGATCCATTTTCAAAATTTT 47
1 CCACAGGCTTGATTGAAATAAATTTGATCCATTTTCAAAATTTT 47

```

```

RESULT 2
AF257499/c LOCUS AF257499 93153 bp DNA linear HTG 26-JUL-2002
DEFINITION Homo sapiens chromosome 8 clone GSI-179c23 map 8p23.1, WORKING
DRAFT SEQUENCE, 7 unordered pieces.
AF257499
ACCESSION AF257499.3 GI:14327872
VERSION AF257499.3 GI:14327872
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_CANCELLED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 93153)
Lehmann, R., Menzel, U., Polley, A., Reichwald, K., Schudy, A.,
Sidiqul, R., Taudien, S., Wen, G., Schlegelberger, B., Siebert, R.,
Rosenthal, A. and Platzner, M.
Chromosome 8 genomic sequence
Unpublished
2 (bases 1 to 93153)
Schudy, A., Schilhabel, M., Baumgart, C., Menzel, U., Schattevoy, R. and
Rosenthal, A.
Direct Submission
Submitted (18-Apr-2000) Genome Analysis, Institute of Molecular
Biotechnology, Bartenbergstrasse 11, Jena 07745, Germany
On Jun 7, 2001 this sequence version replaced gi:8151509.
----- Genome Center
Center: Institute of Molecular Biotechnology
Center code: IMB
Web site: http://genome.imb-jena.de/

```

Contact: gscj-submit@genome.imb-jena.de

----- Project Information

Center Project name: H127

Center Clone name: GSI-179c23

----- Summary Statistics

Sequencing vector: M13; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 84498 bases at least Q40

Consensus quality: 89860 bases at least Q30

Quality coverage: 5,12 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 7 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

1 6411: contig of 6411 bp in length
* 6412 6511: gap of unknown length
* 6512 52674: contig of 46163 bp in length
* 52675 52774: gap of unknown length
* 52775 66702: contig of 13928 bp in length
* 66703 66802: gap of unknown length
* 66803 73537: contig of 6835 bp in length
* 73538 73737: gap of unknown length
* 73738 79228: contig of 5491 bp in length
* 79229 79328: gap of unknown length
* 79329 82459: contig of 3131 bp in length
* 82460 82559: gap of unknown length
* 82560 93153: contig of 10594 bp in length.

```

FEATURES

source

```

1..93153
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="8"
/map="8p23.1"
/clone="GSI-179c23"

```

ORIGIN

```

Query Match 99.1%; Score 46.6; DB 2; Length 93153;
Best Local Similarity 97.9%; Pred. No. 5.5e-05;
Matches 46; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
1 CCACAGGCTTGATTGAAATAAATTTGATCCATTTTCAAAATTTT 47
1 CCACAGGCTTGATTGAAATAAATTTGATCCATTTTCAAAATTTT 56583

```

```

RESULT 3
AC011835/c LOCUS AC011835 182151 bp DNA linear PRI 11-DEC-2001
DEFINITION Homo sapiens chromosome 8, clone RP11-143D15, complete sequence.
ACCESSION AC011835
VERSION AC011835.8 GI:17488682
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 182151)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Unpublished
2 (bases 1 to 182151)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baldwin, J., Barna, N., Beckerly, R., Boguslavsky, L., Boukhgalter, B.,
Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,
Cooke, P., Dearlano, K., Dewar, K., Domingo, M., Donelan, L., Doyle, M.,

```

Pereira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heatford, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, T., Lehoczyk, J., Lieu, C., Locke, K., Macdonald, P., Margolis, N., McEwan, P., McGuirk, A., McKernan, K., McLaughlin, J., Meltrim, J., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testafaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

TITLE

JOURNAL

REFERENCE

AUTHORS

Submitted (15-OCT-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 182151)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouckgalter, B., Brown, A., Camarata, J., Campio, A., Chang, J., Chazaro, B., Choquel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R., Landers, T., Lehoczyk, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Margolis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nduku, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Riese, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testafaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zaitoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

JOURNAL

REFERENCE

AUTHORS

Submitted (04-SEP-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 182151)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouckgalter, B., Brown, A., Camarata, J., Campio, A., Chang, J., Chazaro, B., Choquel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R., Landers, T., Lehoczyk, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Margolis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nduku, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Riese, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testafaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zaitoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

JOURNAL

REFERENCE

AUTHORS

Submitted (11-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Dec 11, 2001 this sequence version replaced gi:1542061.

ALL repeats were identified using RepeatMasker:

Smit, A. F. A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

FEATURES

source

Contact: sequence submissions@genome.wi.mit.edu
----- Project Information
Center project name: 12584
Center clone name: 143_D_15

Location/Qualifiers

source

1. 182151
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-133D15"
/clone_lib="RP11-11 Human Male BAC"
complement(512..597)
/rpt_family="Char11e1a"
complement(610..913)
/rpt_family="Char11e1a"
2178..2258
/rpt_family="MIR3"
2276..2406
/rpt_family="MIR"
4083..4241
/rpt_family="MIR3"
5608..5891
/rpt_family="AluSx"
6464..6775
/rpt_family="AluY"
6910..7023
/rpt_family="L2"
8176..8197
/rpt_family="AT_rich"
9419..9485
/rpt_family="MIR"
10035..10435
/rpt_family="MER47B"
complement(11642..11929)
/rpt_family="AluSq"
12215..12529
/rpt_family="AluSx"
12900..13193
/rpt_family="AluSx"
14043..14393
/rpt_family="TRA)n"
14405..14503
/rpt_family="TRA)n"
14532..14965
/rpt_family="TRA)n"
complement(15222..15800)
/rpt_family="Tigger7"
15675..15679
/note="30 qual SINGL region"
complement(15801..16106)
/rpt_family="AluY"
15827..15836
/note="30 qual SINGL region"
complement(16107..16215)
/rpt_family="Tigger7"
18288..18598
/rpt_family="AluSx"
20572..20598
/rpt_family="AT_rich"
20830..20858
/rpt_family="CAT)n"
complement(21041..21334)
/rpt_family="AluSx"
complement(22065..22181)
/rpt_family="L1M4"
complement(22365..22545)
/rpt_family="L1MC1"
complement(22630..22944)
/rpt_family="AluSx"
23695..24210

```

repeat_region /rpt_family="MERIA"
24726..24818 /rpt_family="MIR"
repeat_region complement(25222..25545)
/rpt_family="LIMBS"
repeat_region complement(25660..25726)
/rpt_family="MIR"
repeat_region 25905..26202 /rpt_family="AluSx"
repeat_region 27319..27425 /rpt_family="AluSx"
repeat_region 27426..27522 /rpt_family="(TTAA)n"
repeat_region 27559..28668 /rpt_family="AT_rich"
repeat_region 28750..28775 /rpt_family="L1PA5"
repeat_region /rpt_family="(A)n"
23067..29212 /rpt_family="AT_rich"
repeat_region 30630..31560 /rpt_family="L2"
repeat_region 31593..31621 /rpt_family="AT_rich"
repeat_region 33828..33887 /rpt_family="(GA)n"
repeat_region complement(34814..35105) /rpt_family="AluSg"
repeat_region complement(35165..36274) /rpt_family="L1ME1"
repeat_region complement(36275..36531) /rpt_family="AluSx"
repeat_region complement(36532..37050) /rpt_family="L1ME1"
repeat_region complement(37063..37228) /rpt_family="L1ME1"
repeat_region complement(37715..37801)

```

Query Match 99.1%; Score 46.6; DB 9; Length 182151;
 Best Local Similarity 97.9%; Pred. No. 4.9e-05;
 Matches 46; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCACGAGCTGTATGAAATTAATTTGATCACCATTTCGAATTTT 47
 Db 31661 CCACGAGCTGTATGAAATTAATTTGATCACCATTTCGAATTTT 31615

RESULT 4
 LOCUS AC016065/c 185463 bp DNA linear PRI 01-DEC-2001
 DEFINITION Homo sapiens chromosome 8, clone RP11-115C21, complete sequence.
 ACCESSION AC016065
 VERSION AC016065.14 GI:17223348
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 185463)
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 Homo sapiens chromosome 8, clone RP11-115C21
 Unpublished
 2 (bases 1 to 185463)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
 Baldwin, J., Barna, N., Beckerly, R., Boguslavsky, L., Boukhgalter, B.,
 Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,
 Cooke, P., Deatellano, K., Dewar, K., Domino, M., Donegan, L., Doyle, M.,
 Ferrelia, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,
 Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
 Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
 Lehoczy, J., Lieu, C., Locke, K., Macdonald, P., Margulis, N.,
 McEwan, P., McGurk, A., McKernan, K., McLaughlin, U., Meltrim, J.,
 Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
 Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,

Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talmas, J.,
 Testafaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
 Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (20-NOV-1999) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 185463)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B.,
 Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
 Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
 Cooke, P., Deatellano, K., Dewar, K., Diaz, J. S., Dodge, S., Fero, S.,
 Ferrelia, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
 Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K.,
 Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
 Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J.,
 Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
 Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, U., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnpack, R.,
 Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N.,
 Struss, N., Subramanian, A., Talamas, J., Testafaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
 Zaitoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
 JOURNAL
 COMMENT
 Direct Submission
 Submitted (01-DEC-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Dec 1, 2001 this sequence version replaced gi:16924135.
 All repeats were identified using RepeatMasker:
 Smt, A F. A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center project name: 14199
 Center clone name: 115_C_21

 FEATURES
 source
 1..185463
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="8"
 /map="8"
 /clone="RP11-115C21"
 /clone_11b="RPCT-11 Human Male BAC"
 29..326
 /rpt_family="AluSx"
 repeat_region 1443..1549 /rpt_family="(TTAA)n"
 repeat_region 1550..1646 /rpt_family="AT_rich"
 repeat_region 1683..2792 /rpt_family="AT_rich"
 /rpt_family="L1PA5"
 1722..1729
 /note="<30 qual SNGL region"
 2874..2898
 /rpt_family="(A)n"
 repeat_region 3190..3263 /rpt_family="AT_rich"
 repeat_region 3268..3333 /rpt_family="(TTA)n"
 repeat_region 4743..5540 /rpt_family="L2"
 repeat_region 5712..5740

```

repeat_region /rpt_family="AT_rich"
7947. .8004
repeat_region /rpt_family="(GA)n"
complement(8931. .9222)
repeat_region /rpt_family="AluSg"
complement(9243. .10391)
repeat_region /rpt_family="L1ME1"
complement(10392. .10648)
repeat_region /rpt_family="AluSx"
complement(10649. .11167)
repeat_region /rpt_family="L1ME1"
complement(11180. .11345)
repeat_region /rpt_family="L1ME1"
complement(11832. .11937)
repeat_region /rpt_family="Trigger7"
complement(11931. .12426)
repeat_region /rpt_family="Trigger7"
13849. .14153
repeat_region /rpt_family="AluSg"
14626. .14878
repeat_region /rpt_family="AluY"
complement(14984. .15284)
repeat_region /rpt_family="AluJo"
15303. .15386
repeat_region /rpt_family="MIR3"
15495. .15540
repeat_region /rpt_family="A-rich"
17156. .17230
repeat_region /rpt_family="AT_rich"
complement(17808. .18120)
repeat_region /rpt_family="AluSg"
complement(18195. .19089)
repeat_region /rpt_family="MERS4B"
19123. .19232
repeat_region /rpt_family="MER31-int"
19292. .19418
repeat_region /rpt_family="MER31-int"
complement(19421. .19549)
repeat_region /rpt_family="FLAM_C"
complement(19727. .19793)
repeat_region /rpt_family="Alu"
complement(19832. .19882)
repeat_region /rpt_family="Charle1"
20637. .20946
repeat_region /rpt_family="AluY"
21160. .21499
repeat_region /rpt_family="MER31-int"
21500. .21807
repeat_region /rpt_family="AluSx"
21806. .21913
repeat_region /rpt_family="MER31-int"
22546. .23228
repeat_region /rpt_family="MER31-int"
23683. .23989
repeat_region /rpt_family="AluSx"
complement(25070. .25360)
repeat_region /rpt_family="AluJo"
26370. .26675
repeat_region /rpt_family="AluJo"
27392. .27528
repeat_region /rpt_family="AluSx"
27525. .27660
repeat_region /rpt_family="AluSx"
27661. .27963
repeat_region /rpt_family="AluSx"
27964. .28294
repeat_region /rpt_family="AluSx"
28295. .28350
repeat_region /rpt_family="(TGGGA)n"
complement(28719. .28842)
repeat_region /rpt_family="MLT1L"
complement(29249. .29551)
repeat_region /rpt_family="AluSx"

```

```

repeat_region complement(31417. .31677)
repeat_region /rpt_family="MIR"
complement(31868. .32002)
repeat_region /rpt_family="L2"
32628. .32941
repeat_region /rpt_family="AluSx"
33756. .33887
repeat_region /rpt_family="FLAM_C"
33905. .34179
repeat_region /rpt_family="L1ME1"
34838. .34991
repeat_region /rpt_family="AluSg/x"
35321. .35630
repeat_region /rpt_family="AluSg"
complement(35769. .36076)
repeat_region /rpt_family="AluY"
36214. .36385
repeat_region /rpt_family="MERSB"
complement(36646. .36956)
repeat_region /rpt_family="AluSx"
complement(40004. .40102)
repeat_region /rpt_family="L2"
40547. .40576
repeat_region /rpt_family="AT_rich"
41201. .41344
repeat_region /rpt_family="L2"
complement(41415. .41750)
repeat_region /rpt_family="MLT1A2"
43260. .43526

Query Match 99.1%; Score 46.6; DB 9; Length 185463;
Best Local Similarity 97.9%; Pred. No. 4.9e-05;
Matches 46; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCACAGCTTGATTGAAATTAASITTGATCACCATTTTCAATTTT 47
Db 5780 CCACAGCTTGATTGAAATTAAGTTGATCACCATTTTCAATTTT 5734

RESULT 5
AC018398 191377 bp DNA linear PRI 30-APR-2002
LOCUS Homo sapiens chromosome 8, clone RP11-16G12, complete sequence.
DEFINITION AC018398
ACCESSION AC018398
VERSION AC018398.10 GI:20198704
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 191377)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 8, clone RP11-16G12
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 191377)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
Boguslavsky,L., Boukhgalter,B., Brown,A., Castle,A., Colangelo,M.,
Collins,S., Collymore,A., Cooke,P., Dearlano,P., Dewar,K.,
Domino,M., Doyle,M., Fenesor,J., Ferreira,P., FitzHugh,M.,
Forrest,C., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B.,
Heaford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L.,
Karacas,A., Klein,J., Landers,T., Lehoczy,J., Lieu,C., Locke,K.,
Macdonald,P., Margus,N., McEwan,P., McGurk,A., McKernan,K.,
Meldrim,J., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T.,
O'Donnell,P., Peterson,K., Pierre,N., Pollara,V., Riley,R.,
Rothman,D., Roy,A., Santos,R., Severy,P., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
TITLE Direct Subdivision
JOURNAL Submitted (10-DEC-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

```

REFERENCE
AUTHORS

3 (bases 1 to 191377)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Bouhagalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., Dearrellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,
Landers, T., Lehoczeky, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., Meltrim, J., Menus, L.,
Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliwer, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retetta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, U., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Straus, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
Direct Submission
Submitted (19-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 191377)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Bouhagalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., Dearrellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,
Landers, T., Lehoczeky, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., Meltrim, J., Menus, L.,
Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliwer, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retetta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, U., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Straus, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
Direct Submission
Submitted (30-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 19, 2002 this sequence version replaced g1:201481178.
All repeats were identified using RepeatMasker:
Smt, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RW/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: I3535
Center clone name: 16_G_12

FEATURES
source
----- Location/Qualifiers
1. 191377
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-16G12"

repeat_region
/rpt_family="AluSg/x"
complement(1277, .1428)
/rpt_family="L1ME1"
repeat_region
1732, .1823
/rpt_family="CA)n"
2184, .2238
/note="PCR product sequence only"
complement(2649, .2774)
/rpt_family="L2"
repeat_region
2963, .3133
/rpt_family="TG)n"
3167, .3208
/rpt_family="TG)n"
3581, .3587
/note="<30 qual SNGL region"
complement(6732, .6840)
/rpt_family="L2"
complement(8405, .8540)
/rpt_family="MIR3"
8922, .9027
/rpt_family="L2"
9061, .9174
/rpt_family="L2"
9390, .9595
/rpt_family="L2"
complement(9592, .9699)
/rpt_family="MIR"
11091, .11333
/rpt_family="AluSx"
11359, .11649
/rpt_family="AluSp"
11765, .12100
/rpt_family="AluDb"
12325, .12593
/rpt_family="AluDb"
13535, .13813
/rpt_family="AluSp"
13817, .13850
/rpt_family="CMAA)n"
13857, .14036
/rpt_family="MIR"
17230, .17345
/rpt_family="L1ME1"
17528, .17551
/rpt_family="AT_rich"
18471, .18757
/rpt_family="AluDo"
21255, .21353
/rpt_family="L2"
21420, .21484
/rpt_family="L1PA9"
22294, .22456
/rpt_family="MIR"
complement(12452, .22513)
/rpt_family="MIR"
22665, .23024
/rpt_family="MT1A1"
23390, .23520
/rpt_family="MER85"
23538, .23739
/rpt_family="MER2"
23908, .24013
/rpt_family="MIR"
24043, .24073
/rpt_family="TTTTG)n"
complement(24079, .24356)
/rpt_family="AluSg"
24877, .24952
/rpt_family="LTR67"
27866, .28203
/rpt_family="AluSx"

```
repeat_region      28205..28272
                    /rpt_family="GA-rich"
repeat_region      28663..28682
                    /rpt_family="(T)n"
repeat_region      complement(28688..29140)
                    /rpt_family="MLTIC"
repeat_region      29551..29760
                    /rpt_family="L2"
repeat_region      complement(29782..30036)
                    /rpt_family="LIMEd"
repeat_region      30200..30591
                    /rpt_family="MLTIB"
repeat_region      complement(30615..30910)
                    /rpt_family="LIMEd"
repeat_region      31381..31634
                    /rpt_family="L2"
repeat_region      31679..31699
                    /rpt_family="AT-rich"
repeat_region      31939..31990
                    /rpt_family="L2"
repeat_region      33132..33161
                    /rpt_family="AT-rich"

Query Match      99.1%; Score 46.6; DB 9; Length 191377;
Best Local Similarity 97.9%; Pred. No. 4.9e-05;
Matches 46; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CCACAGCGTTGATTAGAAATTAATTTGATTCACCATTTCAAAATTT 47
Db      116244 CCACAGCGTTGATTAGAAATTAATTTGATTCACCATTTCAAAATTT 116198

RESULT 6
LOCUS      AX087869      240825 bp      DNA      linear      PAT 17-MAR-2001
DEFINITION      Sequence 1 from Patent WO0114550.
ACCESSION      AX087869
VERSION      AX087869.1 GI:13396862
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE      1
AUTHORS      Barry,C., Chumakov,I. and Blumenfeld,M.
TITLE      Prostate cancer-related gene 3 (pg3) and biallelic markers thereof
JOURNAL      Patent: WO 0114550-A 1 01-MAR-2001;
GENSET      (PR)
FEATURES
    source      1..240825
                Location/Qualifiers
                1..240825
                /organism="Homo sapiens"
                /mol_type="unassigned DNA"
                /db_xref="taxon:9606"
                1..2000
                /note="5;regulatory region"
                1823..1840
                /note="5-390.pu"
                1980..1998
                /note="5-390-177.mis"
                1987..2011
                /bound_molecy="5-390-177.probe"
                1999
                /note="5-390-177 : polymorphic base G or C"
                2000..2018
                /note="5-390-177.mis complement"
                2001..2079
                /note="exon A"
                2108..2125
                /note="5-390.rp complement"
                4559..4577
                /note="5-391.pu"
                4582..4600
                /note="5-391-43.mis"

    misc_feature      1823..1840
                /note="5;regulatory region"
    primer_bind      1823..1840
                /note="5-390.pu"
    primer_bind      1980..1998
                /note="5-390-177.mis"
    primer_bind      1987..2011
                /note="5-390-177.mis"
    misc_binding      1987..2011
                /bound_molecy="5-390-177.probe"
    variation      1999
                /note="5-390-177 : polymorphic base G or C"
    primer_bind      2000..2018
                /note="5-390-177.mis complement"
    exon      2001..2079
                /note="exon A"
    primer_bind      2108..2125
                /note="5-390.rp complement"
    primer_bind      4559..4577
                /note="5-391.pu"
    primer_bind      4582..4600
                /note="5-391-43.mis"

    misc_binding      4589..4613
                /bound_molecy="5-391-43.probe"
    variation      4601
                /note="5-391-43 : polymorphic base A or G"
    primer_bind      4602..4620
                /note="5-391-43.mis complement"
    exon      4627..4718
                /note="exon B"
    primer_bind      4891..4908
                /note="5-391.rp complement"
    primer_bind      10007..10025
                /note="5-392.pu"
    exon      10115..10233
                /note="exon C"
    primer_bind      10209..10227
                /note="5-392-222.mis"
    misc_binding      10216..10240
                /bound_molecy="5-392-222.probe"
    variation      10228
                /note="5-392-222 : polymorphic base G or T"
    primer_bind      10229..10247
                /note="5-392-222.mis complement"
    primer_bind      10267..10285
                /note="5-392-280.mis"
    misc_binding      10274..10298
                /bound_molecy="5-392-280.probe"
    variation      10286
                /note="5-392-280 : polymorphic base G or T"
    primer_bind      10287..10305
                /note="5-392-280.mis complement"
    variation      10370
                /note="5-392-364 : insertion of G"
    primer_bind      10411..10430
                /note="5-392.rp complement"
    exon      26810..26897
                /note="exon D"
    exon      31357..31471
                /note="exon E"
    exon      34261..34404
                /note="exon F"
    exon      37377..37466
                /note="exon S"
    primer_bind      39556..39574
                /note="4-59.rp"
    exon      39704..40858
                /note="exon T"
    primer_bind      39877..39896
                /note="4-58.rp"
    primer_bind      39925..39943
                /note="4-58-318.mis"
    misc_binding      39932..39956
                /bound_molecy="4-58-318.probe"
    variation      39944
                /note="4-58-318 : polymorphic base G or T"
    primer_bind      39945..39963
                /note="4-58-318.mis complement"
    primer_bind      39953..39970
                /note="4-59.pu complement"
    primer_bind      39954..39972
                /note="4-58-289.mis"
    misc_binding      39961..39985
                /bound_molecy="4-58-289.probe"
    variation      39973
                /note="4-58-289 : polymorphic base G or C"
    primer_bind      39974..39992
                /note="4-58-289.mis complement"
    primer_bind      40242..40259
                /note="4-58.pu complement"
    primer_bind      41137..41154
                /note="4-54.rp"
    primer_bind      41366..41384
                /note="4-54-199.mis"
    misc_binding      41373..41397
```

```
primer_bind /bound_moiety="4-54-199.probe"
41385..41403
/ote="4-54-180.mis"
variation 41385
/ote="4-54-199 : polymorphic base A or C"
primer_bind 41386..41404
/ote="4-54-199.mis complement"
misc_binding 41392..41416
/ound_moiety="4-54-180.probe"
41404
variation /ote="4-54-180 : polymorphic base A or C"
41405..41423
/ote="4-54-180.mis complement"
primer_bind 41564..41581
/ote="4-54.pu complement"
42122..42141
/ote="4-51.rp"
primer_bind 42213..42231
/ote="4-51-312.mis"
42220..42244
misc_binding /ound_moiety="4-51-312.probe"
42232
variation /ote="4-51-312 : polymorphic base G or C"
42233..42251
/ote="4-51-312.mis complement"
primer_bind 42526..42543
/ote="4-51.pu complement"
50436..50545
exon /ote="exon G"
67289..67309
primer_bind /ote="99-86.rp"
67456..67474
/ote="99-86-266.mis"
67463..67487
misc_binding /ound_moiety="99-86-266.probe"
67475
variation /ote="99-86-266 : polymorphic base A or G"
67476..67494
primer_bind /ote="99-86-266.mis complement"
67724..67741
primer_bind /ote="99-86.pu complement"
63182..69200
/ote="4-88.rp"
primer_bind 69502..69520
/ote="4-88-107.mis"
69509..69533
misc_binding /ound_moiety="4-88-107.probe"
69521
variation /ote="4-88-107 : polymorphic base A or G"
69522..69540
primer_bind /ote="4-88-107.mis complement"
69609..69626
primer_bind /ote="4-88.pu complement"
72698..72715
/ote="5-397.pu"
primer_bind 72819..72837
/ote="5-397-141.mis"
72826..72850
misc_binding /ound_moiety="5-397-141.probe"
72838
variation /ote="5-397-141 : polymorphic base G or T"
72839..72857
primer_bind /ote="5-397-141.mis complement"
72881..72918
exon /ote="exon H"
73099..73117
/ote="5-397.rp complement"
75858..75877
primer_bind /ote="5-398.pu"
75989..76151
exon /ote="exon I"
76041..76059
primer_bind /ote="5-398-203.mis"
```

```
misc_binding 76048..76072
/ound_moiety="5-398-203.probe"
variation 76060
/ote="5-398-203 : polymorphic base A or C"
primer_bind 76061..76079
/ote="5-398-203.mis complement"
primer_bind 76289..76306
/ote="5-398.rp complement"
81006..81025
/ote="99-12738.pu"
primer_bind 81234..81252
/ote="99-12738-248.mis"
misc_binding 81241..81265
/ound_moiety="99-12738-248.probe"
81253
variation 81253
Query Match 99.1%; Score 46.6; DB 6; Length 240825;
Best Local Similarity 97.9%; Pred.No. 4.7e-05;
Matches 46; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCACAGGCTTGATAGAAATAAATTTGATCCACATTTCAATTTT 47
Db 83764 CCACAGGCTTGATAGAAATAAATTTGATCCACATTTCAATTTT 83810
RESULT 7
AX523960 240825 bp DNA linear PAT 21-NOV-2002
LOCUS Sequence 1 from Patent WO02066641.
ACCESSION AX523960
VERSION AX523960.1 GI:25168889
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS Barry,C. and Chumakov,I.
TITLE Pg-3 and diallelic markers thereof
JOURNAL Patent: WO 02066641-A 1 29-AUG-2002;
GENSET (FR)
FEATURES
Location/Qualifiers
source 1..240825
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1..2000
/ote="5-regulatory region"
primer_bind 1823..1840
/ote="5-390.pu"
primer_bind 1980..1998
/ote="5-390-177.mis"
1987..2011
/ound_moiety="5-390-177.probe"
1999
variation /ote="5-390-177 : polymorphic base G or C"
2000..2018
/ote="5-390-177.mis complement"
2001..2079
exon /ote="exon A"
2108..2125
primer_bind /ote="5-390.rp complement"
4559..4577
/ote="5-391.pu"
primer_bind 4582..4600
/ote="5-391-43.mis"
4589..4613
misc_binding /ound_moiety="5-391-43.probe"
4601
variation /ote="5-391-43 : polymorphic base A or G"
4602..4620
primer_bind /ote="5-391-43.mis complement"
4627..4718
exon
```



```
/note="exon B"
primer_bind 4891..4908
primer_bind /note="5-391.rp complement"
10007..10025
exon /note="5-392.pu"
10115..10233
primer_bind /note="exon C"
10209..10227
misc_binding /note="5-392-222.mis"
10216..10240
variation /bound_motif="5-392-222.probe"
10228
primer_bind /note="5-392-222 : polymorphic base G or T"
10229..10247
primer_bind /note="5-392-222.mis complement"
10267..10285
misc_binding /note="5-392-280.mis"
10274..10298
variation /bound_motif="5-392-280.probe"
10286
primer_bind /note="5-392-280 : polymorphic base G or T"
10287..10305
variation /note="5-392-280.mis complement"
10370
primer_bind /note="5-392-364 : insertion of G"
10411..10430
exon /note="5-392.rp complement"
26810..26897
exon /note="exon D"
31357..31471
exon /note="exon E"
34261..34404
exon /note="exon F"
37377..37466
exon /note="exon S"
39556..39574
primer_bind /note="4-59.rp"
39704..40858
exon /note="exon T"
39877..39896
primer_bind /note="4-58.rp"
39925..39943
primer_bind /note="4-58-318.mis"
39932..39956
misc_binding /bound_motif="4-58-318.probe"
39944
variation /note="4-58-318 : polymorphic base G or T"
39945..39963
primer_bind /note="4-58-318.mis complement"
39953..39970
primer_bind /note="4-59.pu complement"
39954..39972
primer_bind /note="4-58-289.mis"
39961..39985
variation /bound_motif="4-58-289.probe"
39973
primer_bind /note="4-58-289 : polymorphic base G or C"
39974..39992
primer_bind /note="4-58-289.mis complement"
40242..40259
primer_bind /note="4-58.pu complement"
41137..41154
primer_bind /note="4-54.rp"
41366..41384
primer_bind /note="4-54-199.mis"
41373..41397
misc_binding /bound_motif="4-54-199.probe"
41385..41403
variation /note="4-54-180.mis"
41385
primer_bind /note="4-54-199 : polymorphic base A or C"
41386..41404
primer_bind /note="4-54-199.mis complement"
41392..41416
misc_binding /bound_motif="4-54-180.probe"
41404
variation /note="4-54-180 : polymorphic base A or C"
41405..41423
primer_bind /note="4-54-180.mis complement"
41564..41581
primer_bind /note="4-54.pu complement"
42122..42141
primer_bind /note="4-51.rp"
42213..42231
misc_binding /note="4-51-312.mis"
42220..42244
variation /bound_motif="4-51-312.probe"
42232
primer_bind /note="4-51-312 : polymorphic base G or C"
42233..42251
primer_bind /note="4-51-312.mis complement"
42526..42543
primer_bind /note="4-51.pu complement"
50436..50545
exon /note="exon G"
67289..67309
primer_bind /note="99-86.rp"
67456..67474
primer_bind /note="99-86-266.mis"
67463..67487
misc_binding /bound_motif="99-86-266.probe"
67475
variation /note="99-86-266 : polymorphic base A or G"
67476..67494
primer_bind /note="99-86-266.mis complement"
67724..67741
primer_bind /note="99-86.pu complement"
69182..69200
primer_bind /note="4-88.rp"
69502..69520
primer_bind /note="4-88-107.mis"
69509..69533
misc_binding /bound_motif="4-88-107.probe"
69521
variation /note="4-88-107 : polymorphic base A or G"
69522..69540
primer_bind /note="4-88-107.mis complement"
69609..69626
primer_bind /note="4-88.pu complement"
72698..72715
primer_bind /note="5-397.pu"
72819..72837
misc_binding /note="5-397-141.mis"
72826..72850
variation /bound_motif="5-397-141.probe"
72838
exon /note="5-397-141 : polymorphic base G or T"
72839..72857
primer_bind /note="5-397-141.mis complement"
72881..72918
primer_bind /note="exon H"
73099..73117
primer_bind /note="5-397.rp complement"
75858..75877
exon /note="5-398.pu"
75989..76151
primer_bind /note="exon I"
76041..76059
misc_binding /note="5-398-203.mis"
76048..76072
variation /bound_motif="5-398-203.probe"
76060
primer_bind /note="5-398-203 : polymorphic base A or C"
76061..76079
primer_bind /note="5-398-203.mis complement"
76289..76306
```

primer_bind /note="5-398.rp complement"
81006. .81025
primer_bind /note="99-12738.pu"
81234. .81252
misc_binding /note="99-12738-248.mis"
81241. .81265
/bound_molecly="99-12738-248.probe"
variation 81253

Query Match 99.1%; Score 46.6; DB 6; Length 240825;
Best Local Similarity 97.9%; Pred. No. 4.7e-05;
Matches 46; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCACAGCTTGATTGAAATAAATTGATCACCATTTCATAATTTT 47
|||||
Db 83764 CCACAGCTTGATTGAAATAAATTGATCACCATTTCATAATTTT 83810

RESULT 8
HS591B8/c 142552 bp DNA linear PRI 04-MAR-2003
LOCUS Human DNA sequence from clone RP4-591B8 on chromosome 1p13.1,
DEFINITION complete sequence.
ACCESSION AL035410
VERSION AL035410.7 GI:4775617
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 142552)
Whitehead, S.

AUTHORS Direct Submission
TITLE Submitted (04-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
JOURNAL On May 11, 1999 this sequence version replaced gi:4678848.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

COMMENT

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormep This sequence was generated from part of bacterial clone configs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr1

RP4-591B8 is from the library RP4-4 constructed by the group of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2.

FEATURES

SOURCE
1. 142552
Location/Qualifiers
/organism="Homo sapiens"

/mol_type="genomic DNA"
/db_xref="R2PD:RP4-591B8" /db_xref="taxon:9606"
/chromosome="1"
/map="p13.1"
/clone="RP4-591B8"
/clone_11b="RP4-4"

Query Match 62.6%; Score 29.4; DB 9; Length 142552;
Best Local Similarity 80.5%; Pred. No. 14;
Matches 33; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 6 GGCTTGATTGAAATAAATTGATCACCATTTCATAATTTT 46
|||||
Db 87755 GGCTTGATTGAAATAAATTGATCACCATTTCATAATTTT 87715

RESULT 9
AC025631/c 160530 bp DNA linear HTG 10-JUN-2000
LOCUS Home sapiens clone RP11-2104, WORKING DRAFT SEQUENCE, 4 unordered
DEFINITION pieces.
ACCESSION AC025631
VERSION AC025631.3 GI:8439937
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 160530)
Birren, B., Linton, L., Nussbaum, C. and Lander, E.

AUTHORS Unpublished
TITLE 2 (bases 1 to 160530)
JOURNAL
REFERENCE
AUTHORS Birren, B., Linton, L., Nussbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bedalov, F., Boguslavsky, L., Bourkhalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Grand-pierre, N., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Labocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, C., Liu, G., Locke, K., MacDonald, P., Margulis, N., McCarthy, M., McEwan, P., McGuirk, A., McKernan, K., McPheters, R., Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N., Pisanl, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Sudramanian, A., Talamas, J., Teste, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission
TITLE Submitted (12-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Jun 10, 2000 this sequence version replaced gi:7960380. All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Web site: http://www-seq.wi.mit.edu

Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information

Center project name: I6305
Center clone name: 21_O4
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads

JOURNAL Submitted (17-NOV-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
COMMENT requests: clonerequest@sanger.ac.uk
On Nov 20, 2000 this sequence version replaced gi:11225749.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: EMBL; SW; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone configs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6
RP11-53121 is from the library RP11-11.3 constructed at the Roswell Park Cancer Institute by the group of Peter de Jong. For further details see http://bacpac.med.buffalo.edu/VECTOR:PBAC3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-53121 it may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.
The true left end of clone RP11-93K22 is at 40380 in this sequence. The true right end of clone RP11-801118 is at 100 in this sequence.

FEATURES

source

1..40479
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP11-53121"
/clone_11b="RP11-11.3"
1..1376
/note="L1PA13 repeat: matches 4799. .6155 of consensus"
repeat_region
1395..2170
/note="L1PA2 repeat: matches 1. .776 of consensus"
repeat_region
2166..7397
/note="L1PA2 repeat: matches 900. .6146 of consensus"
repeat_region
7398..10895
/note="L1MA4A repeat: matches 1471. .5001 of consensus"
repeat_region
10904..11118
/note="L1 repeat: matches 3684. .3909 of consensus"
repeat_region
11162..11393
/note="L1MA4A repeat: matches 5006. .5239 of consensus"
repeat_region
11386..12562
/note="L1MA4A repeat: matches 5122. .6300 of consensus"
repeat_region
14367..14708
/note="MLT1J repeat: matches 1. .351 of consensus"
repeat_region
15785..16013
/note="MIR repeat: matches 16. .247 of consensus"
repeat_region
16956..16999
/note="22 copies 2 mer tt 75% conserved"
repeat_region
18599..18729
/note="MIR repeat: matches 19. .153 of consensus"
repeat_region
19169..19766
/note="L2 repeat: matches 2060. .2672 of consensus"
repeat_region
21174..21354
/note="MER53 repeat: matches 4. .189 of consensus"
repeat_region
21765..21983
/note="AluSg/x repeat: matches 84. .305 of consensus"
repeat_region
21984..22286
/note="AluJo repeat: matches 14. .312 of consensus"
repeat_region
23259..23408

ORIGIN

repeat_region /note="MIR repeat: matches 57. .214 of consensus"
repeat_region 24290..29248
repeat_region /note="L1PA5 repeat: matches 1177. .6141 of consensus"
repeat_region 29249..29535
repeat_region /note="MLT1J repeat: matches 104. .393 of consensus"
repeat_region 30752..31055
repeat_region /note="AluX repeat: matches 5. .295 of consensus"
repeat_region 31090..31115
repeat_region /note="13 copies 2 mer ac 92% conserved"
repeat_region 31166..31299
repeat_region /note="67 copies 2 mer tg 79% conserved"
repeat_region 32267..32575
repeat_region /note="AluY repeat: matches 1. .306 of consensus"
repeat_region 32643..32928
repeat_region /note="L1P repeat: matches 2952. .3240 of consensus"
repeat_region 32929..33366
repeat_region /note="L1PA8 repeat: matches 5727. .6162 of consensus"
repeat_region 33583..33885
repeat_region /note="AluX repeat: matches 1. .311 of consensus"
repeat_region 34284..35039
repeat_region /note="L2 repeat: matches 896. .1674 of consensus"
repeat_region 35523..35817
repeat_region /note="AluX repeat: matches 1. .291 of consensus"
repeat_region 35997..36035
repeat_region /note="L2 repeat: matches 1701. .1737 of consensus"
repeat_region 36036..36372
repeat_region /note="MER7A repeat: matches 1. .346 of consensus"
repeat_region 36373..36938
repeat_region /note="L2 repeat: matches 1737. .2336 of consensus"
repeat_region 37240..37509
repeat_region /note="AluDb repeat: matches 1. .293 of consensus"
repeat_region 38616..39051
repeat_region /note="L2 repeat: matches 2248. .2750 of consensus"

Query Match 58.7%; Score 27.6; DB 9; Length 40479;
Best Local Similarity 75.0%; Pred. No. 67;
Matches 33; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 4 CAGGCTGATTGAAATTAAGTTGATCACCATTTCATTTT 47
Db 13120 CAGGCTGTTGAAATTAAGACTTATTTTCATTTCAATTTCT 13163

RESULT 12
AC090635
LOCUS Homo sapiens chromosome 8 clone RP11-297E18 map 8, LOW-PASS
DEFINITION
SEQUENCE SAMPLING.
AC090635
AC090635.1 GI:13194959
VERSION
HTG; HTGS PHASE0.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eulacerta; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 70766)
REFERENCE
Birken,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 8, clone RP11-297E18
Unpublished
2 (bases 1 to 70766)
Birken,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barra,N., Bastien,V., Boguslavsky,L., Boukhalter,B., Brown,A.,
Cammarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., DeRellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Fato,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J.,
Gardina,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heathord,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T.,
Lehoczky,J., Levine,R., Liu,G., Maclean,C., MacDonald,P.,
Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
McPheters,R., Meldrum,J., Meneus,L., Milnova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,

TITLE
JOURNAL
COMMENT

O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R.,
Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M.,
Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S., Severy, P.,
Sougnuez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talama, J., Testaye, S., Theodore, J.,
Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (04-MAR-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information -----
Center project name: L10552
Center clone name: 297_E_18

----- NOTE: This record contains 85 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1	732: contig of 732 bp in length
733	832: gap of 100 bp
833	1612: contig of 780 bp in length
1613	1712: gap of 100 bp
1713	2443: contig of 731 bp in length
2444	2543: gap of 100 bp
2544	3268: contig of 725 bp in length
3269	3368: gap of 100 bp
3369	4108: contig of 740 bp in length
4109	4208: gap of 100 bp
4209	4938: contig of 730 bp in length
4939	5038: gap of 100 bp
5039	5743: contig of 705 bp in length
5744	5843: gap of 100 bp
5844	6560: contig of 717 bp in length
6561	6660: gap of 100 bp
6661	7398: contig of 738 bp in length
7399	7498: gap of 100 bp
7499	8235: contig of 737 bp in length
8236	8335: gap of 100 bp
8336	9082: contig of 747 bp in length
9083	9182: gap of 100 bp
9183	9923: contig of 741 bp in length
9924	10023: gap of 100 bp
10024	10750: contig of 727 bp in length
10751	10850: gap of 100 bp
10851	11583: contig of 733 bp in length
11584	11683: gap of 100 bp
11684	12415: contig of 732 bp in length
12416	12515: gap of 100 bp
12516	13237: contig of 722 bp in length
13237	13337: gap of 100 bp
13338	14070: contig of 733 bp in length
13338	14071: gap of 100 bp
14071	14892: contig of 722 bp in length
14892	14992: gap of 100 bp
14993	15713: contig of 721 bp in length
15714	15813: gap of 100 bp

15814	16541: contig of 728 bp in length
16542	16641: gap of 100 bp
16642	17389: contig of 748 bp in length
17390	17489: gap of 100 bp
17490	18235: contig of 746 bp in length
18236	18335: gap of 100 bp
18336	19092: contig of 757 bp in length
19093	19192: gap of 100 bp
19193	19936: contig of 744 bp in length
19937	20036: gap of 100 bp
20037	20782: contig of 746 bp in length
20783	20882: gap of 100 bp
20883	21624: contig of 742 bp in length
21625	21724: gap of 100 bp
21725	22454: contig of 730 bp in length
22455	22554: gap of 100 bp
22555	23284: contig of 730 bp in length
23285	23384: gap of 100 bp
23385	24117: contig of 733 bp in length
24118	24217: gap of 100 bp
24218	24937: contig of 720 bp in length
24938	25037: gap of 100 bp
25038	25755: contig of 718 bp in length
25756	25855: gap of 100 bp
25856	26605: contig of 750 bp in length
26606	26705: gap of 100 bp
26706	27434: contig of 723 bp in length
27435	27534: gap of 100 bp
27535	28264: contig of 730 bp in length
28265	28364: gap of 100 bp
28365	29094: contig of 730 bp in length
29095	29194: gap of 100 bp
29195	29944: contig of 750 bp in length
29945	30044: gap of 100 bp
30045	30780: contig of 736 bp in length
30781	30880: gap of 100 bp
30881	31605: contig of 725 bp in length
31606	31705: gap of 100 bp
31706	32424: contig of 719 bp in length
32425	32524: gap of 100 bp
32525	33279: contig of 755 bp in length
33280	33379: gap of 100 bp
33380	34108: contig of 729 bp in length
34109	34208: gap of 100 bp
34209	34944: contig of 736 bp in length
34945	35044: gap of 100 bp
35045	35802: contig of 758 bp in length
35803	35902: gap of 100 bp
35903	36642: contig of 740 bp in length
36643	36742: gap of 100 bp
36743	37473: contig of 731 bp in length
37474	37573: gap of 100 bp
37574	38315: contig of 742 bp in length
38316	38415: gap of 100 bp
38416	39131: contig of 716 bp in length
39132	39231: gap of 100 bp
39233	39964: contig of 733 bp in length
39965	40064: gap of 100 bp
40065	40795: contig of 731 bp in length
40796	40895: gap of 100 bp
40896	41618: contig of 723 bp in length
41619	41718: gap of 100 bp
41719	42457: contig of 719 bp in length
42438	42537: gap of 100 bp
42538	43265: contig of 728 bp in length
43266	43365: gap of 100 bp
43366	44094: contig of 729 bp in length
44095	44194: gap of 100 bp
44195	44928: contig of 734 bp in length
44929	45028: gap of 100 bp
45029	45754: contig of 726 bp in length
45755	45854: gap of 100 bp
45855	46588: contig of 734 bp in length

*	46589	46688: gap of 100 bp
*	46689	47423: contig of 735 bp in length
*	47424	47523: gap of 100 bp
*	47524	48244: contig of 721 bp in length
*	48245	48344: gap of 100 bp
*	48345	49070: contig of 726 bp in length
*	49071	49170: gap of 100 bp
*	49171	49901: contig of 731 bp in length
*	49902	50001: gap of 100 bp
*	50002	50733: contig of 732 bp in length
*	50734	50833: gap of 100 bp
*	50834	51549: contig of 716 bp in length
*	51550	51649: gap of 100 bp
*	51650	52393: contig of 744 bp in length
*	52394	52493: gap of 100 bp
*	52494	53232: contig of 739 bp in length
*	53233	53332: gap of 100 bp
*	53333	54082: contig of 750 bp in length
*	54083	54182: gap of 100 bp
*	54183	54921: contig of 739 bp in length
*	54922	55021: gap of 100 bp
*	55022	55760: contig of 739 bp in length
*	55761	55860: gap of 100 bp
*	55861	56584: contig of 724 bp in length
*	56585	56684: gap of 100 bp
*	56685	57409: contig of 725 bp in length
*	57410	57509: gap of 100 bp

Query Match 58.7%; Score 27.6; DB 2; Length 70766;
Best Local Similarity 75.0%; Pred. No. 61;
Matches 33; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

Oy 4 CAGCGTGTAGTAATAAATTTGATCACCATTTCATTTT 47
Db 52913 CAGCGTGTGTGAATAAGACTTATTTCATTTCAATTC 52956

RESULT 13
AC006193 118335 bp DNA linear PLN 28-MAY-2000
LOCUS Arabidopsis thaliana chromosome I BAC F13011 genomic sequence.
DEFINITION complete sequence.
ACCESSION AC006193 GI:4733953
VERSION AC006193.3
KEYWORDS HTG.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM

REFERENCE AUTHORS
Arabidopsis thaliana (thale cress)
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 118335)
Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
Gonzalez,A., Kremetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
Luros,S., Schwartz,J., Shim,P., Toriumi,M., Vyotskaia,V.S.,
Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.

JOURNAL REFERENCE AUTHORS
Unpublished
2 (bases 1 to 118335)
Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
Gonzalez,A., Kremetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
Luros,S., Schwartz,J., Shim,P., Toriumi,M., Vyotskaia,V.,
Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.

TITLE JOURNAL
Direct Submission
Submitted (09-DEC-1998) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA

REFERENCE AUTHORS
3 (bases 1 to 118335)
Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
Gonzalez,A., Kremetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
Luros,S., Schwartz,J., Shim,P., Toriumi,M., Vyotskaia,V.,
Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.

TITLE Direct Submission
JOURNAL Submitted (04-MAY-1999) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
REFERENCE AUTHORS
4 (bases 1 to 118335)
Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
Gonzalez,A., Kremetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
Luros,S., Schwartz,J., Shim,P., Toriumi,M., Vyotskaia,V.,
Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
TITTLE JOURNAL
Direct Submission
Submitted (11-JUN-1999) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
REFERENCE AUTHORS
5 (bases 1 to 118335)
Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
Altafi,H., Nguyen,M., Lam,B., Southwick,A., Ecker,J., Theologis,A.
and Davis,R.W.
TITTLE JOURNAL
Direct Submission
Submitted (28-MAY-2000) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
COMMENT On May 4, 1999 this sequence version replaced gi:4678192.
FEATURES
source
1. 118335
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="1"
/map="g1147"
/clone="F13011"
1092..2311
/gene="F13011.1"
join(1092..1225,1474..2311)
/gene="F13011.1"
/note="Hypothetical Protein"
/codon_start=1
/protein_id="AAD38245.1"
/db_xref="GI:5042406"
/translation="MGQLIDPKCTNNTONEEFESTKLSAFOYTAQOITVMKIQ
RPTKMKDKMKSSEPTMTTRIPDPHSTGVREDADSVCKFWKLTENPYCSQS
QOHQORAKFIWDLNFIVFMESELSKQADEIKELKALDYERKARRAEIMTKAK
DVEERMRARAEAMQNRKLFKELSEKSEMYMKADLEERQMRRLAEVLRERVQK
IMDARFLBEKLSLESEANROGERERNRMMPKILERACSPARRCEPQIKRGINP
PPRVRAIRSKSEKSKXLEOKYQVLKILLQKTPRCPLPLSSPP"
complement(3352..3963)
/gene="F13011.2"
complement(3352..3963)
/gene="F13011.2"
/note="Hypothetical Protein"
/codon_start=1
/protein_id="AAD38246.1"
/db_xref="GI:5042407"
/translation="WGNCILFGGLDGEEDLLIKVKSQGVLEFPYPTAGVSHGS
GHALFSAVDLMLKPLANDHLVPGQSYVLPFIVSDELKTFGVSHVSNSSLSAIT
PYRSLDGNHRVLRKVTDFVSRRNSHRIKROEKTRRRRTSSKGAIMKVMILINTER
LIOLSLDEGRTNELLESVRAVAKETSITSSSENFLSVQT"
8107..9734
/gene="F13011.3"
join(8107..8140,8222..8358,8440..8489,8562..8893,
8959..9062,9144..9219,9304..9734)
/gene="F13011.3"
/note="very similar to alcohol dehydrogenase; similar to
gi|1143388 class III ADH, glutathione-dependent
formaldehyde dehydrogenase. [Arabidopsis thaliana];
location of est 94M15T7 (gi|2757072) and 94M15XP (3')
(gi|3450565)"
/codon_start=1
/protein_id="AAD38247.1"
/db_xref="GI:5042408"
/translation="METGKVIYTCAAVAMGAGEPLVEDYKVDPPQRLVRIILFT
SICHTDLSAMKGENDORAYPIIGHEAAGIVESGVEEMMAGDHVLPITGCGD

CHVCKRDKANLCERFRVDPMKVMTVDGKTRFPTSNDKPKIYHFLNTSTFSEXTVIDS
ACVLKVPDPEPLEKISLSCGVSQDCKLWMEGIVGAMVNDIOPASTVAFPGAGVGL
AVABARGASAKTIGIDINPDKEQFAGISSEFINKPSKADVAHERMEITGEVGE
YAFECAGSIEALREARFLSTNSGCVATVMGVHASPOLLPIHHEMELFQGSITRASVCG
FKPKTQLPFFITQGLNLNDFISHLPPFDHINEMQOLHOGKALRCLLHLH"
/complement(10308..11958)
/gene="F13011.4"
/complement(10308..10729,10802..10883,11042..11125,
11208..11384,11461..11617,11723..11958))
/gene="F13011.4"
/note="Location of ests BEH3T7 (gi|2414136), 62B10T7
(gi|91264), 22B15T7A (gi|1217083).. 89% identical to
membrane related protein CP5 [Arabidopsis thaliana]"
/codon_start=1
/product="Putative membrane related protein"
/protein_id="AAD38248.1"
/db_xref="GI:5042409"
/translation="MMGLSILVFTFHSIGFSDLIIVFVAPLMAVAVGVGVMWRPK
MAYLDSKLSNSPKPLNQLPISILKTSNHHVPSLSKEKGTGFTVDDEFRHLMKVE
VVDGSPCWIMMDRSTPFSYQAMRDPENGPQVRSRTVPEDATPEWVRDPEWDEF
RKMDMLXSTLERKDTGTMYVQWAKFPFECSDREYIIGRIINDAGRPFYCTIK
GQYPSVPQNKPRRYDLYSSWCLRAVSKSDGEMTCEVLFHEDMGJLWEIAK
LGVROGMGAVAKKIEPLRAYQAKAGLSPSALMAHINTKVAEEFMNBRGSIAE
VTGDKPTGKNIPKILVVGGAIALACTLDKGLTKAVIFGVARFARMGRM"
13535..15717
/gene="F13011.5"
/join(13535..13647,13808..14022,14535..14734,14807..15315,
15402..15717)
/gene="F13011.5"
/note="Identical to Tubulin Alpha-1 Chain [Arabidopsis
thaliana] (gi|135391) Location of ests 13920T077
(gi|2762968), 214C14T7 (gi|1159108)."
/codon_start=1
/product="alpha tubulin"
/protein_id="AAD38249.1"
/db_xref="GI:5042410"
/translation="MREIISIHIGAGIOVNSCWLTYCLHEGIDOPDGMPSSTVGA
CHDAPRPSSTSSCOHVPRAVLDLEPVIDEVRGTQRLPHFQQLSGEDANN
FARGHTVREIVDTLERKLDADCTGQGLVNAVGGSLSLRLSLVD
FGRKSLGFTIYSPQVSTAVVEPVNSVLSLHTHTVVLNDEALYDLCRSLD
IRPRTYSLNRLISQTSILTSLSLRDGAINDITFQNLVPPRIHMLSYAVPI
SSAKVHEQPSVEITTSVSEPSNMMAKCDPRHKGACCLMAGVGVPRDVTAVAA
IKAKRTIOPWDCPTGFKGICINQPSVPGDLAVGVAVCMISNTTAVAFSRID
HKFDLMSKRAVHMVYVGEEMGESEARBDLALAEKYBEVSGCAEDDEBGEY
"
/complement(16208..17133)
/gene="F13011.6"
/complement(16208..16338,17040..17133))
/gene="F13011.6"
/note="Unknown protein; Location of ests 247E21T7
(gi|241341)"
/codon_start=1
/protein_id="AAD38250.1"
/db_xref="GI:5042411"
/translation="MAEPRKATAVVXKMDLPEDDDEFEEFEINEDMLEKEVKEYSQ
QMEDMDDDVNDVDSRQLRKELNGTIDK"
/complement(17641..19522)
/gene="F13011.7"
/complement(17641..18251,18334..18833,19281..19522))
/gene="F13011.7"
/note="31% identical to Glucan endo-1,3-beta-glucosidase
precursor (11-33)-beta glucan endohydrolase [Triticum
aestivum] (gi|1706551)"
/codon_start=1
/product="Similar to glucan endo-1,3-beta-glucosidase
precursor"
/protein_id="AAD38251.1"
/db_xref="GI:5042412"
/translation="MATKRLPKTVVQMLKNNINKVKFLDADETTMGALAGGLEVM
VAIPNDQLVMTSDEKADKVRKQVTRVYNDGVNITFVAGSGEPDLKSYGSLFILT
PALANIONALNEAGLNSGVKATVPLNADVDSVPSNPAAGRPDLIGQMTQIVD
FLKNNAPITINIPPLSLYGNDDPFLNVAFPDGAIRPGLINDYNNVDANPDIVS
SLKAVGHGDMPIIVGEVGMPTGDKRANAGSAIRFYNGLLPRLGNTKGPIAPRTYIEV

YLFGLIDEDAKSIAPGPFERNHGIKFKFDQPKFIDLSGQGSKFLIGANVPYLPNK
WCTPFEARADLTKRLANIDYACTCTALGSSSCNTLDANASVAFPMFQVQKQ
DSACAFQGLATITNTQINISQGCNPFIOIVASASSFSOSSSVSLVLIWFLLSGMP
"
20970..22826
/gene="F13011.8"
/join(20970..21383,21459..21893,22659..22826)
/gene="F13011.8"
/note="Hypothetical Protein"
/codon_start=1
/protein_id="AAD38252.1"
/db_xref="GI:5042413"
/translation="MASLISFSLPPKAVRSSISAPOTQITNTERLEDFGRKGIKF
SSNNIPWELKVRNGSSSLKLSIDHVSVPKRYKXGEGFEVLYTVGDSRGV
GVYINDEEPKGVSSVSGSDMSVDTSDADALQIELCTAGVLDITYVSLVPS
MATLVKRNKRPPTLKRGINSYLRFKRSNGIOLGICCSICPPPISSPELLSP
SEAMKAESSGWFGESEGERPGIWAEDSVITTLERKMSIYGAPPAERLKAAYNTPS
KFEITIDQDKMDVWRDVCIPRIDVQVWEATLLRVHRSNVNARSRCSLRDMERSN
GNRA"
/complement(24442..25986)
QY 4 CAGGCTGATTAATAASTTTGATCACCATTTCAATTTT 47
Db 70308 CAGGCTTGATTAAGCAATAGATTGATTTGATTAGAACTTT 70351
Query Match 58.7%; Score 27.6; DB 8; Length 118335;
Best Local Similarity 75.0%; Pred. No. 56;
Matches 33; Conservative 1; Mismatches 10; Indels 0; Gaps 0;
gene
AC108046/c 177483 bp DNA linear PRI 20-MAR-2002
LOCUS AC108046 AC024020
DEFINITION Homo sapiens BAC clone Rpl1-308K2 from 4, complete sequence.
AC108046 GI:19551179
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 177483)
AUTHORS Sultston,J.B. and Waterston,R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 177483)
AUTHORS Levy,A., Kozlowicz,A., Creason,K. and Pearman,C.
TITLE The sequence of Homo sapiens BAC clone Rpl1-308K2
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 177483)
AUTHORS Waterston,R.H.
TITLE Direct Submision
JOURNAL Submitted (24-JAN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 177483)
Waterston,R.
Direct Submision
Submitted (20-MAR-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Mar 20, 2002 this sequence version replaced gi:18642919.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wuicl.edu/gsc
Contact: saplens@watscn.wuicl.edu
----- Summary Statistics
Center project name: H_NH00808K02
Drafting Center: WIBR

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P. Y., Zhao, B., Frengen, E., Tateo, M., Catanesse, J. J. and de Jong, P. J. (1998). An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-696F12; the clone sequenced to the right is RP11-264F11. Actual start of this clone is at base position 1 of RP11-308K2; actual end is at base position 177483 of RP11-308K2.

The sequence of AC024020 has been incorporated into AC108046.

FEATURES

source

Location/Qualifiers

1..177483

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="4"

/map="4"

/clone="RP11-308K2"

/clone_lib="RPCI-11"

1..153

/rpt_family="L1"

154..464

/rpt_family="Alu"

465..566

/rpt_family="L1"

567..873

/rpt_family="Alu"

874..956

/rpt_family="L1"

968..1037

/rpt_family="L1"

1121..1729

/rpt_family="L1"

1863..2148

/rpt_family="MALR"

3785..4013

/rpt_family="CT-rich"

4019..4143

/rpt_family="L1"

4157..4451

/rpt_family="Alu"

6910..8125

/rpt_family="L1"

8171..8290

/rpt_family="MIR"

8416..8489

/rpt_family="L1"

repeat_region

8490..8802

/rpt_family="Alu"

repeat_region

8803..8960

/rpt_family="L1"

repeat_region

8961..9012

/rpt_family="(TA)n"

repeat_region

9013..9112

/rpt_family="L1"

repeat_region

9367..9475

/rpt_family="MIR"

repeat_region

10800..10822

/rpt_family="(TG)n"

repeat_region

10823..11110

/rpt_family="Alu"

repeat_region

12603..12692

/rpt_family="MIR"

repeat_region

12815..12913

/rpt_family="(TATTA)n"

repeat_region

14271..14396

/rpt_family="L2"

repeat_region

15058..15345

/rpt_family="Alu"

repeat_region

15992..16057

/rpt_family="L2"

repeat_region

16412..16853

/rpt_family="L1"

repeat_region

17001..17037

/rpt_family="AT-rich"

repeat_region

20274..20295

/rpt_family="AT-rich"

repeat_region

21690..21805

/rpt_family="MIR"

repeat_region

22201..22228

/rpt_family="AT-rich"

repeat_region

22229..22467

/rpt_family="L1"

repeat_region

22681..22976

/rpt_family="Alu"

repeat_region

23210..23552

/rpt_family="L1"

repeat_region

24520..24549

/rpt_family="AT-rich"

repeat_region

24648..24766

/rpt_family="MER1_type"

repeat_region

24777..24898

/rpt_family="Alu"

repeat_region

24907..25000

/rpt_family="MIR"

repeat_region

28464..28913

/rpt_family="L1"

repeat_region

28941..29039

/rpt_family="MER103"

repeat_region

29831..29897

/rpt_family="L2"

repeat_region

30233..30371

/rpt_family="MALR"

repeat_region

30372..30841

/rpt_family="L1"

repeat_region

30855..31383

/rpt_family="L1"

repeat_region

31391..31779

/rpt_family="L1"

repeat_region

31810..32339

/rpt_family="MALR"

repeat_region

32826..32860

/rpt_family="AT-rich"

repeat_region

35299..35393

/rpt_family="MIR"

repeat_region

35468..35991

/rpt_family="MER2_type"

repeat_region

36009..36168

/rpt_family="MER2_type"

repeat_region

36121..36174


```

repeat_region      /rpt_family="MER2_type"
                    36365..36399
repeat_region      /rpt_family="(CA)n"
                    37093..37119
repeat_region      /rpt_family="(TTTG)n"
                    38314..38591
repeat_region      /rpt_family="L1"
                    39520..39623
repeat_region      /rpt_family="MIR"
                    39992..40066
repeat_region      /rpt_family="MIR"
                    40471..40635
repeat_region      /rpt_family="MIR"

```

```

Query Match      57.9%; Score 27.2; DB 9; Length 177483;
Best Local Similarity 80.0%; Pred. No. 69;
Matches 32; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

Oy      8 CTTGATTAGAAATTAATTTGATCACCATTTCAAATTTT 47
Db      136383 CTTGTTAGAAATTAATTTACTTACCATTTCCTTTT 136344

```

```

RESULT 15
BX469934/c      225586 bp      DNA      linear      VRT 25-SEP-2003
LOCUS      Zebrafish DNA sequence from clone CH211-209P5, complete sequence.
ACCESSION      BX469934
VERSION      BX469934.5      GI:35760789
KEYWORDS      HTG.
SOURCE      Dario rerio (zebrafish)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
      Cypriniformes; Cyprinidae; Danio.
      1 (bases 1 to 225586)
      Wood, J.
      Direct Submission
      Submitted (25-SEP-2003) Wellcome Trust Sanger Institute, Hinxton,
      Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
      zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
      On Sep 25, 2003 this sequence version replaced gi:35209448.
      ----- Genome Center
      Center: Wellcome Trust Sanger Institute
      Center code: SC
      Web site: http://www.sanger.ac.uk
      Contact: zfish-help@sanger.ac.uk
      -----

```

COMMENT

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep

Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhirong Bao and Sean Eddy, submitted), and those

```

beginning 'drr' were identified by Rick Waterman (Stephen Johnson lab, Mashu). For further information see
http://www.sanger.ac.uk/Projects/D\_rerio/fishmask.shtml
is from a CHOR1-211 BAC library
VECTOR: pTABAC2.1.
Location/Qualifiers
1..225586
/organism="Dario rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-209P5"
/clone_1ib="CHOR1-211"

```

ORIGIN

```

Query Match      57.9%; Score 27.2; DB 5; Length 225586;
Best Local Similarity 76.2%; Pred. No. 67;
Matches 32; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

```

```

Oy      5 AGGCTTGATTAGAAATTAATTTGATCACCATTTCAAATTTT 46
Db      70317 AGGATTGACTAGAAATTAATTTGATCACCATTTCAAATTT 70276

```

```

RESULT 16
BX908795/c      237042 bp      DNA      linear      HTG 20-FEB-2004
LOCUS      Dario rerio clone DKEX-92J12, WORKING DRAFT SEQUENCE, 10 unordered
DEFINITION      pieces.
ACCESSION      BX908795
VERSION      BX908795.4      GI:42733234
KEYWORDS      HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE      Dario rerio (zebrafish)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
      Cypriniformes; Cyprinidae; Danio.
      1 (bases 1 to 237042)
      Burton, J.
      Direct Submission
      Submitted (19-FEB-2004) Wellcome Trust Sanger Institute, Hinxton,
      Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
      zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
      On Feb 20, 2004 this sequence version replaced gi:4106643.
      ----- Genome Center
      Center: Wellcome Trust Sanger Institute
      Center code: SC
      Web site: http://www.sanger.ac.uk
      Contact: zfish-help@sanger.ac.uk
      ----- Project Information
      Center project name: ZK92J12
      ----- Summary Statistics
      Assembly program: XGAP4; version 4.5
      Chemistry: Dye-terminator; 100% of reads
      Consensus quality: 232035 bases at least Q40
      Consensus quality: 232682 bases at least Q30
      Consensus quality: 233528 bases at least Q20
      Insert size: 236142; sum-of-contigs
      Insert size: 234948; 2.6% error; agarose-fp
      Quality coverage: 11.47x in Q20 bases; sum-of-contigs Quality
      coverage: 11.69x in Q20 bases; agarose-fp
      -----

```

COMMENT

* NOTE: This is a 'working draft' sequence. It currently consists of 10 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 57482: contig of 57482 bp in length

* 57483 57582: gap of 100 bp

* 57583 109940: contig of 52358 bp in length

* 109941 110040: gap of 100 bp

* 110041 114134: contig of 4094 bp in length
* 114135 114234: gap of 100 bp
* 114235 149477: contig of 35243 bp in length
* 149478 149577: gap of 100 bp
* 149578 164819: contig of 15242 bp in length
* 164820 164919: gap of 100 bp
* 164920 185164: contig of 20245 bp in length
* 185165 202375: contig of 17111 bp in length
* 202376 202475: gap of 100 bp
* 202476 204766: contig of 2291 bp in length
* 204767 204866: gap of 100 bp
* 204867 224261: contig of 19395 bp in length
* 224262 224361: gap of 100 bp
* 224362 237042: contig of 12681 bp in length.
FEATURES
source 1. 237042
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone_idb="DnioxKey"
1. 57482
/note="assembly_fragment:03190
fragment_chain:1"
misc_feature 57583. .109940
/note="assembly_fragment:02229
fragment_chain:1"
misc_feature 110041. .114134
/note="assembly_fragment:00001
fragment_chain:1"
misc_feature 114235. .149477
/note="assembly_fragment:01609
fragment_chain:1"
misc_feature 149578. .164819
/note="assembly_fragment:00308
fragment_chain:1"
misc_feature 164920. .185164
/note="assembly_fragment:01211
fragment_chain:2"
misc_feature 185265. .202375
/note="assembly_fragment:00585
fragment_chain:2"
misc_feature 202476. .204766
/note="assembly_fragment:04241
fragment_chain:2"
misc_feature 204867. .224261
/note="assembly_fragment:00880
fragment_chain:3"
misc_feature 224362. .237042
/note="assembly_fragment:00061
fragment_chain:3"
ORIGIN
Query Match 57.9%; Score 27.2; DB 2; Length 237042;
Best Local Similarity 76.2%; Pred. No. 66;
Matches 32; Conservative 1; Mismatches 9; Indels 0; Gaps 0;
Oy 5 AGGCTTGATTTGAAATTAATTTGATCCATTTTCAATT 46
Db 141139 AGGATGACTAGAAATATGTTGAATCCCATTTTACATT 141098
RESULT 17
AL592222 160796 bp DNA linear ROD 19-SEP-2003
LOCUS AL592222
DEFINITION Mouse DNA sequence from clone RP23-366M19 on chromosome 11,
complete sequence.
ACCESSION AL592222
VERSION AL592222.11 GI:34850734
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Submitted (19-SEP-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Sep 19, 2003 this sequence version replaced gi:27899567.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Emi, EMBL; Swi, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C_elegans/wormpep -----
Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
RP23-366M19 is from the RPCI-23 Mouse BAC library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
Location/Qualifiers
source 1. 160796
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-366M19"
/clone_idb="RPCI-23"
FEATURES
source 1. 160796
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-366M19"
/clone_idb="RPCI-23"
ORIGIN
Query Match 56.2%; Score 26.4; DB 10; Length 160796;
Best Local Similarity 71.7%; Pred. No. 13e+02;
Matches 33; Conservative 1; Mismatches 12; Indels 0; Gaps 0;
Oy 1 CCACAGGCTTGATTTGAAATTAATTTGATCCATTTTCAATT 46
Db 94768 CCACAGGCTTGCTTAAGTGTCTGTAGAGCCATTTTCAATT 94723
RESULT 18
BX321875 169994 bp DNA linear VRT 01-SEP-2004
LOCUS BX321875
DEFINITION zebrafish DNA sequence from clone CH211-254C8 in linkage group 16,
complete sequence.
ACCESSION BX321875
VERSION BX321875.7 GI:50724671
KEYWORDS HTG.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

REFERENCE 1 (bases 1 to 169994)
AUTHORS Cypriniformes; Cyprinidae; Danio.
TITLE Pelan.S.
JOURNAL Direct Submission
Submitted (01-SEP-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk
On Jul 27, 2004 this sequence version replaced gi:12398497.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: zfish-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep
Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.
Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhirong Bao and Sean Eddy, submitted), and those beginning 'dir' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see http://www.sanger.ac.uk/Projects/D_rexio/fishmask.shtml
is from a CHORI-211 BAC library
VECTOR: pTRABAC2.1.
Location/Qualifiers
1..169994
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-254C8"
/clone_11b="CHORI-211"
ORIGIN
Query Match 56.2%; Score 26.4; DB 5; Length 169994;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 33; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
Qy 3 ACAGGCTTGATTGAATAAAGTTGATCACCATTTCAAATTT 46
Db 17472 ACATTTTATTTATTTATTAATCTGTCATCATTTTCAAAAT 17515
RESULT 19
BX936440 197773 bp DNA linear HTG 17-FEB-2004
LOCUS BX936440
DEFINITION Danio rerio clone DKEY-1K24, *** SEQUENCING IN PROGRESS ***, 9
uncovered pieces.
ACCESSION BX936440
VERSION BX936440
KEYWORDS GI:42592599
SOURCE HTG; HTGS_PHASE1.
ORGANISM Danio rerio (zebrafish)
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 197773)
AUTHORS Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
TITLE Cypriniformes; Cyprinidae; Danio.
JOURNAL Direct Submission
Submitted (15-FEB-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk
On Feb 17, 2004 this sequence version replaced gi:42557958.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: ZK1K24
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 195294 bases at least Q40
Consensus quality: 195829 bases at least Q30
Consensus quality: 196182 bases at least Q20
Insert size: 196973; sum-of-contigs
Insert size: 186171; 1.5% error; agarose-fp
Quality coverage: 6.96x in Q20 bases; sum-of-contigs Quality coverage: 7.36x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently consists of 9 contigs. The true order of the pieces * is not known and their order in this sequence record is * arbitrary. Gaps between the contigs are represented as * runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence * as soon as it is available and the accession number will * be preserved.
1 7723: contig of 7723 bp in length
* 7724 7823: gap of 100 bp
* 7824 43716: contig of 35893 bp in length
* 43717 43816: gap of 100 bp
* 43817 97564: contig of 53748 bp in length
* 97565 97664: gap of 100 bp
* 97665 127663: contig of 29999 bp in length
* 127664 127763: gap of 100 bp
* 127764 137218: contig of 9455 bp in length
* 137219 137318: gap of 100 bp
* 137319 157224: contig of 19906 bp in length
* 157225 157324: gap of 100 bp
* 157325 168221: contig of 10897 bp in length
* 168222 168321: gap of 100 bp
* 168322 174805: contig of 6484 bp in length
* 174806 174905: gap of 100 bp
* 174906 197773: contig of 22868 bp in length.
Location/Qualifiers
1..197773
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-1K24"
/clone_11b="DanioKey"
1..7723
/note="assembly fragment:00077
fragment_chain:1"
7824..43716
/note="assembly fragment:00859
fragment_chain:1"
43817..97564
/note="assembly fragment:01613
fragment_chain:1"
97665..127663
/note="assembly fragment:01223
fragment_chain:1"
127764..137218
/note="assembly fragment:00163
fragment_chain:1"

```

misc_feature      fragment chain:1"
137319.._157224
/note="assembly_fragment:00366
fragment_chain:1"
157325.._168231
/note="assembly_fragment:00260
fragment_chain:1"
168322.._174805
/note="assembly_fragment:00016
fragment_chain:1"
174906.._197773
/note="assembly_fragment:00603.0"

misc_feature
misc_feature
misc_feature

Query Match      56.2%; Score 26.4; DB 2; Length 197773;
Best Local Similarity 71.7%; Pred. No. 1.2e+02;
Matches 33; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

Db      170985 CTCAGGCTTCATTAAAAAATNCATCCATCTGCATTTTGAATATT 171030

Cy      2 CACAGGCTTGATTGAATATAAATTTGATCCACATTTTCAATTTT 47
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

AC023173      203451 bp      DNA      linear      ROD 20-MAR-2002
LOCUS      Mus musculus chromosome 6 clone RP23-11G22 strain C57BL6/J,
DEFINITION      complete sequence.
ACCESSION      AC023173
VERSION      AC023173.3 GI:19551124
KEYWORDS      HTG.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 203451)
Akhter,N., Antoneilis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C.,
Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E., Laric,P.,
Lee-Jin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
Maguiles,E.H., Masiello,C., Mastrii,B., Mastrii,S.D.,
McCluskey,J.C., McDowell,J., Paguirigan,C., Pearson,R.,
Potnyou,M.E., Prasad,A., Schuler,M.G., Stentrop,S., Thomas,J.W.,
Thomas,P.J., Touchman,J.W., Tsurgou,C., Vogt,J.L., Walker,M.A.,
Weberby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 203451)
Green,E.D.
Direct Submission
Submitted (09-FEB-2000) NIH Intramural Sequencing Center, 8717
Groveomt Circle, Gaitersburg, MD 20877, USA
3 (bases 1 to 203451)
Green,E.D.
Direct Submission
Submitted (20-MAR-2002) NIH Intramural Sequencing Center, 8717
Groveomt Circle, Gaitersburg, MD 20877, USA
On Mar 20, 2002 this sequence version replaced gi:12658031.

COMMENT
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.mouse@hgri.nih.gov
----- Project Information
Center project name: Ya
Center clone name: 011G22

```

all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

CLONE LENGTH: This sequence represents the entire insert of this clone unless otherwise noted. If there are overlapping clones, the overlaps are noted in the beginning and end of the Features section.

Location/Qualifiers

1..203451

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="C57BL6/J"

/db_xref="taxon:10090"

/chromosome="6"

/clone="RP23-11G22"

/clone_1db="RPC1 mouse BAC library 23"

45177..45242

/note="single clone coverage"

49880..49881

/note="low quality, single stranded/single chemistry region"

122920..122971

/note="unresolved tandem repeat; consensus does not reflect repeat; region sized by PCR"

123299..123345

/note="single clone coverage"

124835..124848

/note="single clone coverage"

155484..155530

/note="single clone coverage"

171106..171130

/note="single clone coverage"

195018..203451

/note="clone overlaps with GenBank Accession Number AC024950 (nucleotides 1-8433) clone RP23-208N9 (center project name yb)"

ORIGIN

Query Match 56.2%; Score 26.4; DB 10; Length 203451;
Best Local Similarity 78.9%; Pred. No. 1.2e+02;
Matches 30; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

0y 10 TGATTGAAATAASTTTGATCACCATTTCGAATTT 47
||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 111115 TTATTTAAATAAGTTTGATTCACAATTTTAATTTT 111152

RESULT 21
AC122835/c 218449 bp DNA linear ROD 08-NOV-2003
LOCUS
DEFINITION Mus musculus BAC clone RP23-187118 from 3, complete sequence.
AC122835
VERSION AC122835.3 GI:23592205
KEYWORDS
HTG.
MUS musculus (house mouse)
SOURCE
ORGANISM
MUS musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Wang,C., Bielicki,L., Spalding,L. and Mangiapanello,L.
TITLE The sequence of Mus musculus BAC clone RP23-187118
JOURNAL Unpublished (2001)
REFERENCE
AUTHORS Wilson,R.
TITLE Sequencing of Mus musculus
JOURNAL Unpublished (2001)
REFERENCE
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (25-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE
4 (bases 1 to 218449)

```

AUTHORS      McPherson,J.D. and Waterston,R.H.
TITLE        Direct Submission
JOURNAL      Submitted (01-SEP-2002) Genome Sequencing Center, 4444 Forest Park
              Parkway, St. Louis, MO 63108, USA
REFERENCE    5 (bases 1 to 218449)
AUTHORS      McPherson,J.D. and Waterston,R.H.
TITLE        Direct Submission
JOURNAL      Submitted (08-OCT-2002) Genome Sequencing Center, 4444 Forest Park
              Parkway, St. Louis, MO 63108, USA
REFERENCE    6 (bases 1 to 218449)
AUTHORS      Wilson,R.
TITLE        Direct Submission
COMMENT      Submitted (08-NOV-2003) Department of Genetics, Washington
              University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
              On Oct 8, 2002 this sequence version replaced g1:22475871.
              -----
              Genome Center
              Center: Washington University Genome Sequencing Center
              Center code: WUGSC
              Web site: http://genome.wustl.edu
              Contact: submissions@watson.wustl.edu
              -----
              Summary Statistics
              Center project name: M_BA0187118
              -----

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
http://genome.wustl.edu

SOURCE INFORMATION:
The RNCI-23 BAC Library has been constructed by Kazutoyo Oseegawa
and Minako Tateo in the laboratory of Pieter de Jong
(http://www.chori.org) from female C57BL/6J mouse kidney and/or
brain genomic DNA. The clone and detailed information can be
obtained from Research Genetics, Inc. (http://www.resgen.com) or
Pieter de Jong and coworkers at http://www.chori.org

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone.
Location/Qualifiers
1..21849
 /organism="Mus musculus"
 /mol_type="Genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="3"
 /map="3"
 /clone="RP23-187118"
 /clone_id="RPCI-23"
 4921..5047
 /rpt_family="MIR"
 8201..8301
 /rpt_family="L1"
 12177..12293
 /rpt_family="Alu"
 12842..13147
 /rpt_family="L1"
 13165..13736
 /rpt_family="L1"
 14002..14198
 repeat_region

```

repeat_region	/rpt_family="B2"	15410.15496
repeat_region	/rpt_family="L2"	16050.16250
repeat_region	/rpt_family="B4"	16316.16844
repeat_region	/rpt_family="L1"	17935.18099
repeat_region	/rpt_family="B4"	19256.19343
repeat_region	/rpt_family="Alu"	20833.21026
repeat_region	/rpt_family="B2"	23277.23353
repeat_region	/rpt_family="Alu"	24094.24183
repeat_region	/rpt_family="B4"	25171.25354
repeat_region	/rpt_family="B2"	26413.26495
unsure	/note="Sequence derived from one plasmid subclone."	
repeat_region	/rpt_family="Alu"	27759.27864
repeat_region	/rpt_family="Alu"	30322.30434
repeat_region	/rpt_family="Alu"	30353.30470
repeat_region	/rpt_family="B4"	30478.30685
repeat_region	/rpt_family="B2"	31557.31670
repeat_region	/rpt_family="Alu"	32025.32091
repeat_region	/rpt_family="ERV1"	33011.33137
repeat_region	/rpt_family="L1"	33179.33420
repeat_region	/rpt_family="L1"	33480.33672
repeat_region	/rpt_family="L1"	33678.33761
repeat_region	/rpt_family="L1"	33946.34358
repeat_region	/rpt_family="L1"	35496.35763
repeat_region	/rpt_family="B4"	38602.38729
repeat_region	/rpt_family="Alu"	40746.41002
repeat_region	/rpt_family="B4"	42010.42076
repeat_region	/rpt_family="ID"	44975.45052
repeat_region	/rpt_family="ID"	46218.46395
repeat_region	/rpt_family="MIR"	47806.47918
repeat_region	/rpt_family="B4"	50026.50225
repeat_region	/rpt_family="B2"	53346.53722
repeat_region	/rpt_family="MaLR"	55092.55281
repeat_region	/rpt_family="B2"	55474.55750
repeat_region	/rpt_family="L1"	57416.57796
repeat_region	/rpt_family="MaLR"	59233.59323
repeat_region	/rpt_family="L1"	59877.59999
repeat_region	/rpt_family="B4"	60000.60047
repeat_region	/rpt_family="ERV1"	

```
repeat_region      62085. .62387  
/rpt_family="r1"  
repeat_region      62841. .62930  
/rpt_family="MaR"  
repeat_region      62931. .63326  
/rpt_family="MaR"  
repeat_region      63327. .63549  
/rpt_family="MaR"  
repeat_region      65251. .66094  
/rpt_family="r1"  
repeat_region      66479. .66867  
/rpt_family="MaR"  
repeat_region      67533. .67936  
/rpt_family="MaR"  
repeat_region      68446. .68824  
/rpt_family="r1"  
repeat_region      68829. .69029  
/rpt_family="r1"  
repeat_region      69097. .70572  
/rpt_family="r1"  
repeat_region      70566. .70620  
/rpt_family="r1"  
repeat_region      70834. .70932  
/rpt_family="Alu"  
repeat_region      71286. .71402  
/rpt_family="Alu"  
repeat_region      71468. .71864  
/rpt_family="MaR"  
repeat_region      71994. .72035  
/rpt_family="MaR"  
repeat_region      74029. .74162
```

Query Match	56.2%;	Score 26.4;	DB 10;	Length 218449;
Best Local Similarity	71.7%;	Pred. No. 1.2e+02;		
Matches 33;	Conservative 1;	Mismatches 12;	Indels 0;	Gaps 0;

Qy 2 CACAGGCTTGATTAGAAATTAASTTTGATCCACATTTTCAAATTTT 47
||| ||| : ||| ||| |||
Db 105833 CACAGGGCAGATTCGAATAAATGTGTTTTCAATTTTTT 105788

RESULT 22				
AC119515/c				
LOCUS	AC119515	213417 bp	DNA	linear
DEFINITION	Rattus norvegicus clone CH230-40602, ***		SEQUENCING IN PROGRESS	HTG 15-NOV-2002

ACCESSION AC119515 GI:25012225
VERSION AC119515.5
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus

REFERENCE	1 (bases 1 to 213417)
AUTHORS	Muzny, D. Marie., Metzker, M. Lee., Abramzon, S., Adams, C., Alder, J.,

1 (bases due to 213417)

Muzny, D., Marie, E., Metzker, M., Lee, S., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alshrooke, S., Amin, A., Angiano, D., Anyalabeche, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Bartscheid, M., Bernabed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Day, C., Deay-carroll, L., De Anda, C., Dedetich, D., Delgado, O., Denson, S., Detamco, C., Ding, Y., Dinh, H., Diya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, J., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Gantz, R., Garcia, A., Garner, T., Garza, M., Gebregeregergs, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guetara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harare, Y., Havlik, P., Hawes, A., Henderson, N., Hernandez, J.,

TITLE	Direct Submission
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 213417)
AUTHORS	Worley, K.C.
TITLE	Direct Submission
JOURNAL	Submitted (28-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	3 (bases 1 to 213417)
AUTHORS	Rat Genome Sequencing Consortium.
TITLE	Direct Submission
JOURNAL	Submitted (15-NOV-2002) Human Genome Sequencing Center, Department

COMMENT

at Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 15, 2002 this sequence version replaced gi:23616941.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hpsc.dcm.tmc.edu/projects/rac/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

```

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center Project name: GUDM
Center Clone name: CH230-40602
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 197852 bases at least Q40
Consensus quality: 199838 bases at least Q30
Consensus quality: 201193 bases at least Q20
Estimated insert size: 207854; sum-of-configs estimation
Quality coverage: 5x in Q20 bases; sum-of-configs estimation

```

```

*****
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 213417: contig of 213417 bp in length.
  Location/Qualifiers
    1.213417
      /organism="Rattus norvegicus"
      /mol_type="genomic DNA"
      /db_xref="taxon:10116"
      /clone="CH230-40602"
      1.1798
        /note="wgs_end_extension"
        clone_end="5p6"
        6829..7711
          /note="clone boundary"
          clone_end="5p6"
          site="end_sequence:B2148299"
          end_sequence..208667
            /note="wgs_contig"
            211434..213417
              /note="wgs_contig"

FEATURES
  source
    1.213417
      /organism="Rattus norvegicus"
      /mol_type="genomic DNA"
      /db_xref="taxon:10116"
      /clone="CH230-40602"
      1.1798
        /note="wgs_end_extension"
        clone_end="5p6"
        6829..7711
          /note="clone boundary"
          clone_end="5p6"
          site="end_sequence:B2148299"
          end_sequence..208667
            /note="wgs_contig"
            211434..213417
              /note="wgs_contig"

ORIGIN
  Query Match      55.7%; Score 26.2; DB 2; Length 213417;
  Best Local Similarity 79.5%; Pred. No. 1.4e+02;
  Matches 31; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

  9 TTGATTAGAAATTAATTTGATCACCATTTCATTTTCAATTTT 47
  Db 147592 TTGTTAGAAATTAATTTGCAACATCATTTTTCATTTT 147554

RESULT 23
AC098034/c 217538 bp DNA linear HTG 10-MAY-2003
LOCUS AC098034
DEFINITION Rattus norvegicus clone CH230-194A3, *** SEQUENCING IN PROGRESS
ACCESSION AC098034
VERSION AC098034.5 GI:30521938
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
          1 (bases 1 to 217538)
            Muzny,D.,Matte., Metzker,M.,Lee., Adamzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Bismalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Caesar,H., Center,A.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,K., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Frazer,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,T.,
            Gebregiorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,

```

```

  Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
  Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogue,M.,
  Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
  Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
  Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
  Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
  Liu,Y., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
  Lorusubew,L., Louised,H., Lozada,R.J., Lu,X., Ma,J.,
  Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
  Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
  Manning,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
  Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
  Morgan,M., Morris,K., Morris,S., Mundaas,M., Murphy,M., Nair,L.,
  Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
  Nwaekelemeh,O., Okunnu,G., Olarnunsaogon,A., Pal,S., Parks,K.,
  Pasternak,S., Paul,H., Perez,A., Perez,L., Flannoch,C.,
  Plopper,F., Polidexter,A., Popovic,D., Prims,E., Pu,L.,-L.,
  Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
  Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
  Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
  Sanders,M., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H.,
  Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D.,
  Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Soosa,D.,
  Steimle,M., Strong,R., Sutton,A., Svatek,A., Taber,P., Taylor,C.,
  Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
  Valas,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J.,
  Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
  Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
  Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
  Yu,P., Zhang,J., Zhou,X., Zhou,X., Zhou,S., Dunn,D., von
  Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
  Weinstock,G. and Gibbs,R.A.
  Direct Submission
  Unpublished
  2 (bases 1 to 217538)
  Worley,K.C.
  Direct Submission
  Submitted (23-OCT-2001) Human Genome Sequencing Center, Department
  of Molecular and Human Genetics, Baylor College of Medicine, One
  Baylor Plaza, Houston, TX 77030, USA
  3 (bases 1 to 217538)
  Rat Genome Sequencing Consortium.
  Direct Submission
  Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
  of Molecular and Human Genetics, Baylor College of Medicine, One
  Baylor Plaza, Houston, TX 77030, USA
  On May 10, 2003 this sequence version replaced gi:23096267.
  The sequence in this assembly is a combination of BAC based reads
  and whole genome shotgun sequencing reads assembled using Atlas
  (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
  in the feature table below represents a scaffold in the Atlas
  assembly (a 'contig-scaffold'). Within each contig-scaffold,
  individual sequence contigs are ordered and oriented, and separated
  by sized gaps filled with Ns to the estimated size. The sequence
  may extend beyond the ends of the clone and there may be sequence
  contigs within a contig-scaffold that consist entirely of whole
  genome shotgun sequence reads. Both end sequences and whole genome
  shotgun sequence only contigs will be indicated in the feature
  table.

  ----- Genome Center
  Center: Baylor College of Medicine
  Center code: BCM
  Web site: http://www.hgsc.bcm.tmc.edu/
  Contact: hgsc-help@bcm.tmc.edu
  ----- Project Information
  Center project name: GGAO
  Center clone name: CH230-194A3
  ----- Summary Statistics
  Assembly program: Atlas 3.0?
  Consensus quality: 192804 bases at least Q40
  Consensus quality: 195152 bases at least Q30
  Consensus quality: 197333 bases at least Q20
  Estimated insert size: 201938; sum-of-contigs estimation

```

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 4 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 65820: contig of 65820 bp in length
 * 65821 65920: gap of unknown length
 * 65921 213541: contig of 147621 bp in length
 * 213542 213641: gap of unknown length
 * 213642 215330: contig of 1889 bp in length
 * 215331 215630: gap of unknown length
 * 215631 217538: contig of 1908 bp in length.

Location/Qualifiers
 1. 217538
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-194A3"
 1. 1013
 /note="wgs contig"
 100203. 101969
 /note="wgs_contig"

ORIGIN
 misc_feature
 misc_feature
 misc_feature
 misc_feature

Query Match 55.7%; Score 26.2; DB 2; Length 217538;
 Best Local Similarity 75.6%; Pred. No. 1.4e+02;
 Matches 31; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

Qy 7 GCTTGATTGAAATTAATTTGATCACCATTTCATATTT 47
 Db 48092 GCATGTTGATCATCTTGTATCATCTTTCATATCTT 48052

RESULT 24
 AC107465
 LOCUS AC107465 242655 bp DNA linear HTG 13-MAY-2003
 DEFINITION Rattus norvegicus clone CH230-86H10, *** SEQUENCING IN PROGRESS
 AC107465
 ACCESSION AC107465.6 GI:30580831
 VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 KEYWORDS Rattus norvegicus
 SOURCE Rattus norvegicus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus;
 1 (bases 1 to 242655)
 Muzny,D,Marie, Metzger,M, Lee, A, Abramson, S, Adams, C, Alder, J, Allen, C, Allen, H, Alsbrooks, S, Amin, A, Angilano, D, Anyalebechi, V, Aoyagi, A, Ayodeji, M, Baca, E, Baden, H, Baldwin, D, Bandaranaike, D, Barber, M, Barnstead, M, Benahmed, F, Bismalo, K, Blair, J, Blankenburg, K, Blyth, P, Brown, M, Bryant, N, Buhay, C, Burch, P, Butrell, K, Calderon, E, Cardenas, V, Carter, K, Cavazos, I, Ceasar, H, Center, A, Chacko, J, Chavez, D, Chen, G, Chen, R, Chen, Y, Chen, Z, Chu, J, Cleveland, C, Cockrell, R, Cox, C, Coyle, M, Cree, A, D'Souza, L, Davila, M, L, Davis, C, Davy-Carroll, L, De Anda, C, Dederich, D, Delgado, O, Denson, S, Deramo, C, Ding, Y, Dinh, H, Divya, K, Draper, H, Dugan-Rocha, S, Dunn, A, Durbin, K, Duval, B, Eaves, K, Egan, A, Escotto, M, Eugene, C, Evans, C, A, Falls, T, Fan, G, Fernandez, S, Finley, M, Flagg, N, Forbes, L, Foster, M, Foster, P, Fraser, C, M, Gabisi, A, Ganta, R, Garcia, A, Garner, T, Garza, M, Gebregiorgis, B, Geer, K, Gill, R, Grady, M, Guerra, W, Guevara, W, Gunaratne, P, Haaland, W, Hamil, C, Hamilton, C, Hamilton, K, Harvey, Y, Havlak, P, Hawes, A, Henderson, N, Hernandez, J, Hernandez, R, Hines, S, Hladun, S, L, Hodgson, A, Hogues, M,

Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolyet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovari, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshew, L., Louised, H., Lozdo, R.U., Lu, X., Ma, J., Mangum, B., Mapua, P., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Morgan, M., Morris, K., Morris, S., Mundaas, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K., Nwakoelmen, O., Okwou, G., Olarunpagaon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Prims, E., Pu, L., Puazo, M., Quiroz, J., Rachin, E., Reeves, K., Regier, M.A., Reich, R., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.U., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabors, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tinney, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villalana, D., Walron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, D., Wei, X., White, P., Williams, G., Willson, R., Wlezyk, R., Woden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Gibbs, R.A.

Direct Submission
 Unpublished
 2 (bases 1 to 242655)
 Worley, K.C.

Direct Submission
 Submitted (22-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 242655)
 Rat Genome Sequencing Consortium.

Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On May 13, 2003 this sequence version replaced gi:22855692.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 Project Information
 Center project name: G1XA
 Center clone name: CH230-86H10
 Summary Statistics
 Assembly program: Atlas 3.0;
 Consensus quality: 230159 bases at least Q40
 Consensus quality: 232877 bases at least Q30
 Consensus quality: 234876 bases at least Q20
 Estimated insert size: 245378; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 239139: contig of 239139 bp in length
 * 239140 239239: gap of unknown length
 * 239240 240581: contig of 1342 bp in length
 * 240582 240681: gap of unknown length
 * 240682 242655: contig of 1974 bp in length.

FEATURES
 source
 1. 242655
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-86H10"
 1. 1966
 /note="wgs_contig"
 2365. 5223
 /note="wgs_contig"
 237046. 239139
 /note="wgs_contig"

ORIGIN
 Query Match 55.7%; Score 26.2; DB 2; Length 242655;
 Best Local Similarity 79.5%; Pred. No. 1.4e+02;
 Matches 31; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 9 TTGATTGAATAAATTGATCACCATTTCATTTT 47
 |||
 Db 232179 TTGTTAGAAATTAATTCACATCATTTT 232217

RESULT 25
 AC131879/c
 LOCUS
 DEFINITION
 AC131879 313096 bp DNA linear HTG 15-NOV-2002
 Rattus norvegicus clone CH230-86H14, *** SEQUENCING IN PROGRESS
 *** 4 unordered pieces.
 AC131879.5 GI:25007070
 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 Rattus norvegicus (Norway rat)
 ORGANISM
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 313096)
 Muzny,D.,Marie., Metzker,M.,Lee., Abramson,S., Adams,C., Alder,J.,
 Allen,C., Allen,H., Alpbrooks,S., Amin,A., Anguiano,D.,
 Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
 Balwin,D., Bandaranaike,D., Barber,M., Barnstead,W., Benedmed,F.,
 Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
 Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
 Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A.,
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,J.,
 Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
 Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
 Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
 Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Davies,K.,
 Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
 Frazer,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
 Gebregiorgis,B., Geer,K., Gill,R., Grady,M., Guerra,M., Guevara,W.,
 Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
 Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
 Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
 Hollins,B., Howells,S., Huylk,S., Hume,J., Idlebird,D., Jackson,A.,
 Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,

Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
 Kowis,C., Kraft,C.L., Lebow,H., Lavan,J., Lewis,L., Li,Z., Liu,J.,
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
 Lounsbury,L., Loussed,H., Lozano,R.J., Lu,X., Ma,J.,
 Maheshwari,M., Mahindaratne,M., Mahmoud,M., Mallory,K., Mangum,A.,
 Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,B.,
 Mawlinney,S., McLeod,M.P., McNeill,T.Z., Meenen,B.,
 Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
 Morgan,M., Morris,K., Morris,S., Mundasa,M., Murphy,M., Nair,L.,
 Nankervis,C., Neal,D., Newcom,N., Nguyen,N., Norris,S., Parks,K.,
 Nwankweme,O., Okwodu,G., Olarpunasegoun,A., Pal,S., Parks,K.,
 Pasternak,S., Paul,H., Perez,A., Perez,L., Pfankuch,C.,
 Plapper,F., Pindexter,A., Popovic,D., Prims,E., Pu,L.,
 Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reid,R.,
 Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,R.,
 Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
 Sanders,M., Savary,G., Scherer,S., Scott,G., Shatman,S., Shen,H.,
 Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D.,
 Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Soza,J.,
 Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Uemari,K.,
 Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
 Williams,G., Willson,R., Wleczek,R., Wooden,H., Worley,K.,
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
 Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
 Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
 Weinstein,G., and Gibbs,R.A.

Unpublished
 Direct Submission
 2 (bases 1 to 313096)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (27-AUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 313096)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Nov 15, 2002 this sequence version replaced gi:23689082.

COMMENT
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/atlas/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GSNZ
 Center clone name: CH230-86H14
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 22307 bases at least Q40
 Consensus quality: 225821 bases at least Q30
 Consensus quality: 228136 bases at least Q20
 Estimated insert size: 232385; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 252512: contig of 252512 bp in length
* 252513 252612: gap of unknown length
* 252613 266426: contig of 13814 bp in length
* 266427 266526: gap of unknown length
* 266527 310559: contig of 44033 bp in length
* 310560 310659: gap of unknown length
* 310660 313096: contig of 2437 bp in length.

FEATURES

source

1. .313096
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-86114"
252613. .254069
/note="wgs_contig"

misc_feature

ORIGIN

Query Match 55.7%; Score 26.2; DB 2; Length 313096;
Best Local Similarity 75.6%; Pred. No. 1.3e+02;

Matches 31; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 7 GCTTATTGAAATAAATTGATCAGCATTTTCAAAATTT 47

Db 47852 GCATGGTTGATCATACTTGTATCACCATTTCATATCTT 47812

Search completed: June 2, 2005, 10:20:34
Job time : 1851 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2005, 08:12:23 ; Search time 450 Seconds
(without alignments)
618.284 Million cell updates/sec

Title: US-10-349-143-1

Perfect score: 47
Sequence: 1 ccacagctgtatgaat.....gatacaccatttcaatttc 47

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46.6	99.1	47	3	Aa265654 Human map
2	46.6	99.1	240823	10	Aa265654 Human map
3	46.6	99.1	240825	4	Aa265654 Human map
4	46.6	99.1	240825	6	Aa265654 Human map
5	24.4	51.9	47108	6	ABK1511 Signal tr
6	24.2	51.5	993	9	ADA32589 DNA encod
7	24.2	51.5	4937	6	ABL07356 Drosophi
8	24.2	51.7	5173	6	ABL32951 Human im
9	23.6	50.2	27681	4	Aa336498 Human car
10	23.6	50.2	27681	4	Aa336498 Human car
11	23.6	50.2	27681	4	Aa336498 Human car
12	23.6	50.2	27681	10	AD47191 Human car
13	23.6	50.2	27681	13	AD47191 Human car
14	23.6	50.2	27681	13	AD47191 Human car
15	23.6	50.2	27681	13	AD47191 Human car
16	23.6	50.2	27681	13	AD47191 Human car
17	23.4	49.8	34917	5	Aa336498 Human car
18	23.4	49.8	34917	5	Aa336498 Human car
19	23.2	49.4	564	10	ACF69221 Human pro
20	23.2	49.4	1978	8	ACD13402 Human DNA

21	23.2	49.4	1978	13	ACN41054
22	23.2	49.4	2000	12	AD40809
23	23.2	49.4	2347	11	ADM02167
24	23.2	49.4	3212	4	AB128314
25	23.2	49.4	4368	4	ABK89062
26	23.2	49.4	40359	4	ABJ30382
27	23.2	49.4	61197	11	ACN45130
28	23.2	49.4	73742	12	ADQ97824
29	23.2	49.4	90104	4	AB112402
30	23.2	49.4	90435	12	ADQ59524
31	23.2	49.4	109661	12	ADQ97818
32	23.2	49.4	110000	10	ADQ77343_09
33	23.2	49.4	110000	10	ACF67367_21
34	23.2	49.4	114693	8	Aa265655
35	23.2	49.4	118384	10	AD48308
36	23.2	49.4	126001	12	ABH77123
37	23.2	49.4	243072	10	ACF65382
38	23.2	49.4	360	10	ACF70068
39	23.2	48.9	411	8	ABK54130
40	23.2	48.9	571	6	ABN60688
41	23.2	48.9	3874	4	AB111142
42	23.2	48.9	4254	4	AB111115
43	23.2	48.9	10329	6	AB134122
44	23.2	48.9	12291	4	AAK79265
45	23.2	48.9	12646	4	AB111114
46	23.2	48.9	50000	9	ADH16926
47	23.2	48.9	110000	10	ACF65385_2
48	23.2	48.9	110000	10	ACF67367_31
49	23.2	48.5	1506	2	AAK7506
50	22.8	48.5	2751	2	AAK30434
51	22.8	48.5	2784	2	AAK68270
52	22.8	48.5	28397	12	ADQ59407
53	22.8	48.5	31491	4	AD10203
54	22.8	48.5	31491	4	AAK63301
55	22.8	48.5	31491	6	AAK98863
56	22.6	48.1	360	5	AAK65972
57	22.6	48.1	459	2	AAK75449
58	22.6	48.1	602	8	ACF73906
59	22.6	48.1	701	3	AAK13788
60	22.6	48.1	726	4	AAK94957
61	22.6	48.1	1209	2	AAK83834
62	22.6	48.1	1209	2	AAV53355
63	22.6	48.1	2570	9	ACCS9915
64	22.6	48.1	10174	4	AB103298
65	22.6	48.1	24389	4	AB129006
66	22.6	48.1	28564	10	ADQ48909
67	22.6	48.1	73725	10	ADQ42690
68	22.6	48.1	73725	10	ADQ42690
69	22.6	48.1	73725	10	ADQ42690
70	22.6	48.1	110000	11	ACN44934_1
71	22.6	48.1	245531	13	ABD33022
72	22.6	48.1	322101	10	ADQ58431
73	22.6	48.1	322101	12	AD135046
74	22.6	47.7	472	9	ACH37763
75	22.4	47.7	492	13	ADQ92630
76	22.4	47.7	644	10	ADQ17465
77	22.4	47.7	745	8	ACA04735
78	22.4	47.7	765	4	AAH08429
79	22.4	47.7	2012	12	ADQ20962
80	22.4	47.7	3634	13	ACN42674
81	22.4	47.7	3760	13	ACN42673
82	22.4	47.7	3761	13	ACN42672
83	22.4	47.7	3777	6	AB165195
84	22.4	47.7	3874	4	AB112274
85	22.4	47.7	3886	10	ADQ76730
86	22.4	47.7	4003	12	ADQ24897
87	22.4	47.7	4013	10	ADQ31412
88	22.4	47.7	4032	13	ACN42671
89	22.4	47.7	4688	5	ABV25850
90	22.4	47.7	4688	5	ABV25472
91	22.4	47.7	5085	8	ACF62791
92	22.4	47.7	5085	8	ACF62813
93	22.4	47.7	5087	13	ADQ89577

```
c 94 22.4 47.7 5087 13 ADS89363 Ad889363 Oligonuc
c 95 22.4 47.7 5087 13 ADS89303 Ad889303 Oligonuc
c 96 22.4 47.7 5087 13 ADS89637 Ad889637 Oligonuc
c 97 22.4 47.7 8085 4 Aa866479 Tumour su
c 98 22.4 47.7 8085 6 AbK33986 Human DNA
c 99 22.4 47.7 8085 8 ADA20374 Prostata
c 100 22.4 47.7 8085 8 ADA84181 Human ren
```

ALIGNMENTS

```
RESULT 1
AAZ65654
ID AAZ65654 standard; DNA; 47 BP.
```

```
XX AAZ65654;
AC
XX 11-SEP-2001 (first entry)
DT
XX Human map-related biallelic marker SEQ ID NO.1.
DE
XX Human genome; biallelic marker; high density disequilibrium map;
KM genomic map; haplotype; phenotype; polymorphic base; genotyping;
KW haplotyping; hybridisation; identification; characterisation; diagnosis;
KM single nucleotide polymorphism; SNP; ds.
XX
```

```
OS Homo sapiens.
```

```
XX Key Location/Qualifiers
FH variation repbase(24, C)
FT /tag= a
FT /standard_name= "single nucleotide polymorphism"
```

```
XX WO954500-A2.
XX
XX 28-OCT-1999.
XX
XX 21-APR-1999; 99WO-1B000822.
XX
XX 21-APR-1998; 98US-0082614P.
XX
XX 23-NOV-1998; 98US-0109732P.
XX
```

```
XX (GEST ) GENSET.
XX
XX Cohen D, Blumenfeld M, Chumakov I;
XX
XX WPI; 2000-013267/01.
XX
XX Novel biallelic markers used to construct a high density disequilibrium
XX map of the human genome.
XX
XX Claim 1; Page 230; 2745PP; English.
XX
XX AAZ65654 to AAZ69578 represent human biallelic markers from the present
XX invention, which contain a polymorphic base at position 24 of their
XX nucleotide sequences. AAZ69579 to AAZ77440 represent amplification
XX primers for the biallelic markers. The biallelic markers of the invention
XX have a variety of uses: they can be used for high density mapping of the
XX human genome, and in complex association studies and haplotyping studies
XX which are useful in determining the genetic basis for disease states.
XX Compositions and methods of the invention can also be useful for the
XX identification of the targets for the development of pharmaceutical
XX agents and diagnostic methods, as well as the characterisation of the
XX differential efficacious responses to and side effects from
XX pharmaceutical agents acting on a disease as well as other treatment.
XX N.B. The SEQ ID NOS 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297 and
XX 3367, are not actually given a sequence in the Sequence Listing from the
XX present invention
```

```
XX
XX Sequence 47 BP; 16 A; 8 C; 6 G; 17 T; 0 U; 0 Other;
```

```
Query Match 99.1%; Score 46.6; DB 3; Length 47;
```

```
Best Local Similarity 97.9%; Pred. No. 2.2e-06;
Matches 46; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CCACAGGCTTGATTAATAAATTTGATCACCATTTCAAATTTT 47
Db 1 CCACAGGCTTGATTAATAAATTTGATCACCATTTCAAATTTT 47
```

RESULT 2

```
ADD69391
ID ADD69391 standard; DNA; 240823 BP.
```

```
XX ADD69391;
```

```
XX 15-JAN-2004 (first entry)
```

```
XX Human PG-3 DNA - SEQ ID 127.
```

```
XX fusion protein; fibrinogen-like; coiled-like domain;
```

```
XX angiotensin-related factor; ARF; angiotensin; Ang-1; Ang-2; Ang-3;
```

```
XX Ang-4; Ang-2X; vulnery; antiinflammatory; vasotropic; necrosis;
```

```
XX ischaemia; inflammation; wound healing; CCD; FLD; human; PG-3; ds.
```

```
XX Homo sapiens.
```

```
XX WO2003048185-A2.
```

```
XX 12-JUN-2003.
```

```
XX 21-NOV-2002; 2002WO-US037660.
```

```
XX 30-NOV-2001; 2001US-0334488P.
```

```
XX (GENV-) GENVEC INC.
```

```
XX Kessler PD, Kovsed I;
```

```
XX WPI; 2003-513736/48.
```

```
XX New fusion protein comprising a fibrinogen-like or coiled-like domain,
```

```
XX useful for preparing a composition for treating necrosis, ischaemia or
```

```
XX inflammation, or for promoting wound healing.
```

```
XX Disclosure; SEQ ID NO 127; 340pp; English.
```

```
XX The invention relates to a novel fusion protein comprising a fibrinogen-
```

```
XX like domain (FLD) or coiled-coil domain (CCD). The domain may be
```

```
XX identical or homologous to that of an angiotensin-related factor (ARF),
```

```
XX examples of which include Ang (angiotensin)-1, Ang-2, Ang-3, Ang-4 and
```

```
XX Ang-2X. The molecules of the invention demonstrate vulnery,
```

```
XX antiinflammatory and vasotropic activities whilst the fusion protein may
```

```
XX be useful for preparing a composition for treating necrosis, ischaemia or
```

```
XX inflammation, as well as for promoting wound healing. The current
```

```
XX sequence is that of the human PG-3 DNA of the invention.
```

```
XX Sequence 240823 BP; 66383 A; 50394 C; 51548 G; 72418 T; 0 U; 80 Other;
```

```
Query Match 99.1%; Score 46.6; DB 10; Length 240823;
```

```
Best Local Similarity 97.9%; Pred. No. 5.2e-06;
Matches 46; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CCACAGGCTTGATTAATAAATTTGATCACCATTTCAAATTTT 47
Db 83762 CCACAGGCTTGATTAATAAATTTGATCACCATTTCAAATTTT 83808
```

```
RESULT 3
AAFP24497
ID AAFP24497 standard; cDNA; 240825 BP.
```

```
XX AAFP24497;
XX
XX 23-MAY-2001 (first entry)
```

```
XX Human PG-3 gene.
DE Human; PG-3; cancer; BRCA1; chromosome 8p23; ds.
KM Homo sapiens.
XX
XX Key
FH Location/Qualifiers
FT 1..2000
FT /*tag= a
FT /note= "5' regulatory region"
FT primer_bind
FT 1823..1840
FT /*tag= b
FT primer_bind
FT 1980..1998
FT /*tag= c
FT 1987..2011
FT /*tag= d
FT /note= "binds probe"
FT replace(1999,C)
FT /*tag= e
FT complement(2000..2018)
FT CDS
FT 2001..238825
FT /*tag= f
FT /product= "PG-3"
FT /note= "this sequence contains introns"
FT 2001..2079
FT /*tag= h
FT /label= A
FT primer_bind
FT 2108..2125
FT /*tag= i
FT primer_bind
FT 4559..4577
FT /*tag= j
FT primer_bind
FT 4582..4600
FT /*tag= k
FT primer_bind
FT 4589..4613
FT /*tag= l
FT /note= "binds probe"
FT replace(4601,G)
FT /*tag= m
FT complement(4602..4620)
FT /*tag= n
FT 4627..4718
FT /*tag= o
FT /label= B
FT primer_bind
FT 4691..4908
FT /*tag= p
FT primer_bind
FT 10007..10025
FT /*tag= q
FT exon
FT 10115..10233
FT /label= C
FT /*tag= r
FT primer_bind
FT 10209..10227
FT /*tag= s
FT 10216..10240
FT /*tag= t
FT /note= "binds probe"
FT replace(10228,T)
FT /*tag= u
FT complement(10229..10247)
FT /*tag= v
FT primer_bind
FT 10267..10285
FT /*tag= w
FT 10274..10298
FT /*tag= x
FT /note= "binds probe"
FT replace(10286,T)
FT /*tag= y
FT primer_bind
FT 10287..10305
FT /*tag= z
FT replace(10370,n)
FT /*tag= aa
FT 10411..10430
FT primer_bind

FT exon
FT 26810..26897
FT /*tag= ab
FT /label= D
FT exon
FT 31357..31471
FT /*tag= ad
FT exon
FT 34261..34404
FT /*tag= ae
FT /label= F
FT exon
FT 37377..37466
FT /*tag= af
FT /label= S
FT primer_bind
FT 39556..39574
FT /*tag= ah
FT exon
FT 39704..40858
FT /*tag= ag
FT /label= T
FT primer_bind
FT 39877..39896
FT /*tag= ai
FT primer_bind
FT 39925..39943
FT /*tag= aj
FT misc_binding
FT 39932..39956
FT /*tag= ak
FT /note= "binds probe"
FT replace(39944,T)
FT /*tag= al
FT complement(39945..39963)
FT /*tag= am
FT primer_bind
FT 39953..39970
FT /*tag= an
FT primer_bind
FT 39954..39972
FT /*tag= ao
FT misc_binding
FT 39961..39985
FT /*tag= ap
FT /note= "binds probe"
FT replace(39973,C)
FT /*tag= aq
FT complement(39974..39992)
FT /*tag= ar
FT primer_bind
FT 40242..40259
FT /*tag= as
FT primer_bind
FT 41137..41154
FT /*tag= at
FT primer_bind
FT 41366..41384
FT /*tag= au
FT misc_binding
FT 41373..41397
FT /*tag= av
FT /note= "binds probe"
FT 41385..41403
FT /*tag= ax
FT replace(41385,C)
FT /*tag= aw
FT complement(41386..41404)
FT /*tag= ay
FT 41392..41416
FT /*tag= az
FT /note= "binds probe"
FT replace(41404,C)
FT /*tag= ba
FT complement(41405..41423)
FT /*tag= bb
FT primer_bind
FT 41564..41581
FT /*tag= bc
FT primer_bind
FT 42122..42141
FT /*tag= bd
FT primer_bind
FT 42213..42231
FT /*tag= be
FT 42220..42244
FT /*tag= bf
FT /note= "binds probe"
FT replace(42232,C)
FT /*tag= bg
FT allele
FT primer_bind
FT 10411..10430
FT primer_bind
```

```
FT primer_bind complement(42233..42251)
FT /tag= bh
FT primer_bind 42525..42543
FT /tag= bi
FT exon 50436..50545
FT /tag= bj
FT primer_bind 67289..67309
FT /label= G
FT /tag= bk
FT primer_bind 67456..67474
FT /tag= bl
FT misc_binding 67463..67487
FT /tag= bm
FT /note= "binds probe"
FT replace(67475,G)
FT /tag= bn
FT primer_bind complement(67476..67494)
FT /tag= bo
FT primer_bind 67724..67741
FT /tag= bp
FT primer_bind 69182..69200
FT /tag= bq
FT primer_bind 69502..69520
FT /tag= br
FT misc_binding 69509..69533
FT /tag= bs
FT /note= "binds probe"
FT replace(69521,G)
FT /tag= bt
FT primer_bind complement(69522..69540)
FT /tag= bu
FT primer_bind 69609..69626
FT /tag= bv
FT primer_bind 72698..72715
FT /tag= bw
FT primer_bind 72819..72837
FT /tag= bx
FT misc_binding 72826..72850
FT /tag= by
FT /note= "binds probe"
FT replace(72838,T)
FT /tag= bz
FT primer_bind complement(72839..72857)
FT /tag= ca
FT exon 72881..72918
FT /tag= cb
FT /label= H
FT primer_bind 73099..73117
FT /tag= cc

Query Match 99.1%; Score 46.6; DB 4; Length 240825;
Best Local Similarity 97.9%; Pred.No.5.2e-06;
Matches 46; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCACAGGCTGATTAGAAATTAATTTGATTCACCAATTTCAATTTT 47
DB 83764 CCACAGGCTGATTAGAAATTAATTTGATTCACCAATTTCAATTTT 83810

RESULT 4
ABO81802
ID ABO81802 standard; DNA; 240825 BP.
XX ABO81802;
AC
XX 14-NOV-2002 (first entry)
XX
XX Human PG-3 gene SEQ ID NO:1.
XX
XX Human; PG-3; chromosome 8; chromosome 8p23; polymorphic; SNP;
XX single nucleotide polymorphism; biallelic marker; DNA repair;
XX recombination; cell cycle control; gene; ds.
XX
```

```
OS Homo sapiens.
XX
FH Key
FT 5'UTR
FT primer_bind 1823..1840
FT /tag= a
FT primer_bind 1980..1998
FT /tag= d
FT misc_binding 1987..2011
FT /tag= nh
FT /note= "probe"
FT 1999
FT /tag= ad
FT /standard name= "single nucleotide polymorphism (SNP)"
FT /note= "polymorphic base G or C; the nucleotide is given
FT as an IUPAC ambiguity code in the specification"
FT complement(2000..2018)
FT /tag= dh
FT 2001..2079
FT /tag= b
FT exon 2080..4626
FT /label= A
FT /tag= c
FT /number= 1
FT primer_bind complement(2108..2125)
FT /tag= di
FT 4459..4577
FT /tag= dj
FT primer_bind 4582..4600
FT /tag= dk
FT 4589..4613
FT /tag= ni
FT /note= "probe"
FT 4601
FT /tag= ae
FT /standard name= "single nucleotide polymorphism (SNP)"
FT /note= "polymorphic base A or G; the nucleotide is given
FT as an IUPAC ambiguity code in the specification"
FT complement(4602..4620)
FT /tag= dl
FT 4627..4718
FT /tag= d
FT exon 4719..10114
FT /label= B
FT /tag= e
FT /number= 2
FT primer_bind complement(4891..4908)
FT /tag= dm
FT 10007..10025
FT /tag= dn
FT exon 10115..10233
FT /tag= f
FT /label= C
FT primer_bind 10209..10227
FT /tag= do
FT 10216..10240
FT /tag= nj
FT /note= "probe"
FT 10228
FT /tag= af
FT /standard name= "single nucleotide polymorphism (SNP)"
FT /note= "polymorphic base G or T; the nucleotide is given
FT as an IUPAC ambiguity code in the specification"
FT complement(10229..10247)
FT /tag= dp
FT 10234..26809
FT /tag= g
FT intron
FT /number= 3
FT primer_bind 10267..10285
FT /tag= dq
FT 10274..10298
FT /tag= nk
FT misc_binding
```

```

FT /note= "probe"
FT misc_feature 10286
FT /tag= ag
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /note= "polymorphic base G or T; the nucleotide is given
FT as an IUPAC ambiguity code in the specification"
FT primer_bind /tag= dr
FT complement(10287. .10305)
FT 10370
FT /tag= ah
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /note= "insertion of G; the nucleotide is given as an
FT IUPAC ambiguity code in the specification"
FT primer_bind /tag= ds
FT complement(10411. .10430)
FT 26810. .26897
FT /tag= h
FT /label= D
FT 26898. .31356
FT /tag= i
FT /number= 4
FT 31357. .31471
FT /tag= j
FT /label= B
FT 31472. .34260
FT /tag= k
FT /number= 5
FT 34261. .34404
FT /tag= l
FT /label= F
FT 34405. .37376
FT /tag= m
FT /number= 6
FT 37377. .37466
FT /tag= n
FT /label= S
FT 37467. .39703
FT /tag= o
FT /number= 7
FT 39556. .39574
FT /tag= dc
FT 39704. .40858
FT /tag= p
FT /label= T
FT 39877. .39896
FT /tag= du
FT 39925. .39943
FT /tag= dv
FT 39932. .39956
FT /tag= nl
FT /note= "probe"
FT 39944
FT /tag= ai
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /note= "polymorphic base G or T; the nucleotide is given
FT as an IUPAC ambiguity code in the specification"
FT primer_bind /tag= dw
FT complement(39953. .39970)
FT /tag= dx
FT 39954. .39972
FT /tag= dy
FT 39961. .39985
FT /tag= nm
FT /note= "probe"
FT 39973
FT /tag= aj
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /note= "polymorphic base G or C; the nucleotide is given
FT as an IUPAC ambiguity code in the specification"
FT primer_bind /tag= dz
FT complement(40242. .40259)

```

```

FT /tag= ea
FT intron 40859. .50435
FT /tag= g
FT /number= 8
FT primer_bind 41137. .41154
FT /tag= eb
FT primer_bind 41366. .41384
FT /tag= ec
FT 41373. .41397
FT /tag= nm
FT /note= "probe"
FT primer_bind 41385. .41403
FT /tag= ed
FT 41385
FT /tag= ak
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /note= "polymorphic base A or C; the nucleotide is given
FT as an IUPAC ambiguity code in the specification"
FT misc_feature /tag= ee
FT complement(41386. .41404)
FT 41392. .41416
FT /tag= no
FT /note= "probe"
FT 41404
FT /tag= al
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /note= "polymorphic base A or C; the nucleotide is given
FT as an IUPAC ambiguity code in the specification"
FT primer_bind /tag= ef
FT complement(41405. .41423)
FT 41404
FT /tag= eg
FT primer_bind 42122. .42141
FT /tag= eh
FT primer_bind 42213. .42231
FT /tag= ei
FT 42220. .42244
FT /tag= np
FT /note= "probe"
FT 42232
FT /tag= am
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT misc_feature /tag= am
FT 42232

Query Match 99.1%; Score 46.6; DB 6; Length 240825;
Best Local Similarity 97.9%; Pred. No. 5.2e-06;
Matches 46; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCACAGGCTTGATGAATAAATTAASFTTGATCACCAATTTCAATTTT 47
Db 83764 CCACAGGCTTGATGAATAAATTAAGTTGATCACCAATTTCAATTTT 83810

RESULT 5
ABK31511/c
ID ABK31511 standard; DNA; 47108 BP.
XX
AC ABK31511;
XX
DT 23-APR-2002 (first entry)
XX
DE Signal transduction associated gene modified complementary DNA #177.
XX
KW Human; signal transduction associated gene; cytosine methylation state;
KW Cpg island; signal transduction associated disease; solid tumour; cancer;
KW antitumour; cytostatic; mutant; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200200926-A2.
XX
PD 03-JAN-2002.
XX

```

PF 29-JUN-2001; 2001WO-EP007472.
XX
XX 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIC-) EPIDENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K,
XX
XX WPI; 2002-147896/19.
XX
PT Oligonucleotide for diagnosis and therapy of diseases associated with
PT signal transduction e.g. cancer, comprises chemically modified genomic
PT sequences of genes associated with signal transduction.
XX
XX Claim 1; SEQ ID NO 354; 24pp; English.
XX
CC The present invention relates to chemically modified DNA sequences of
CC signal transduction associated genes. The DNA sequences are chemically
CC modified using a solution of bisulphite, hydrogen sulphite or disulphite.
CC Also disclosed are oligonucleotides and/or PNA oligomers for detecting
CC the cytosine methylation state (CpG islands) of these genes, and a method
CC for the diagnosis and/or therapy of genetic and epigenetic parameters of
CC genes associated with signal transduction. The genomic DNA can be
CC obtained from cells or cellular components which contain DNA, e.g. cell
CC lines, biopsies, blood, sputum, stool, urine, cerebral-spinal fluid,
CC tissue embedded in paraffin such as tissue from eyes, intestine, kidney,
CC brain, heart, prostate, lung, breast or liver, histologic object slides,
CC and all their possible combinations. The sequences of the invention are
CC useful for the diagnosis and therapy of diseases associated with signal
CC transduction e.g. solid tumours and cancer. ABK3158-ABK31545 represent
CC chemically pretreated genomic DNA sequences of different genes associated
CC with signal transduction, or their complementary sequences. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office
XX
SQ Sequence 47108 BP; 15349 A; 429 C; 8736 G; 22594 T; 0 U; 0 Other;
XX
Query Match 51.9%; Score 24.4; DB 6; Length 47108;
Best Local Similarity 77.8%; Pred. No. 1.6e+02;
Matches 28; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
OY 12 ATTGAAATTAASSTTGATCACCATTTCCTCAATTTT 47
DB 25070 ATTAATAAAAACTTAATTAATTAATTTTAAAAATTTT 25035
XX
RESULT 6
ADA32589/c
ID ADA32589 standard; DNA; 993 BP.
XX
XX ADA32589;
XX
DT 20-NOV-2003 (first entry)
XX
DE DNA encoding Acinetobacter baumannii protein #3876.
XX
XX ds; gene; Acinetobacter baumannii; bacterial disease; antibacterial;
XX vaccine; plant biocontrol agent.
XX
XX Acinetobacter baumannii.
XX
XX US6562958-B1.
XX
XX 13-MAY-2003.
XX
XX 04-JUN-1999; 99US-00328352.
XX
XX 09-JUN-1998; 98US-0088701P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX

PI Breton G, Bush D;
XX
XX WPI; 2003-576092/54.
DR P-PSDB; ADA36715.
XX
XX New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
PT for diagnosing a bacterial disease, as components of antibacterial
PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
PT plants.
XX
XX Example; SEQ ID NO 3876; 328pp; English.
XX
XX The invention relates to isolated Acinetobacter baumannii nucleic acids.
CC The A. baumannii nucleic acids and polypeptides are useful as reagents
CC for diagnosing a bacterial disease, as components of antibacterial
CC vaccines, as targets for antibacterial drugs, to detect the presence of
CC A. baumannii and other Acinetobacter species in a sample, in screening
CC compounds for the ability to interfere with the A. baumannii life cycle
CC or to inhibit A. baumannii infection, and as biocontrol agents for
CC plants. The present sequence represents DNA encoding an A. baumannii
CC protein.
XX
SQ Sequence 993 BP; 309 A; 181 C; 203 G; 300 T; 0 U; 0 Other;
XX
Query Match 51.5%; Score 24.2; DB 9; Length 993;
Best Local Similarity 68.1%; Pred. No. 1.3e+02;
Matches 32; Conservative 1; Mismatches 14; Indels 0; Gaps 0;
OY 1 CCACGCGCTGATTGAAATTAASSTTGATCACCATTTCCTCAATTTT 47
DB 205 CAACGATTACTGAGATTAATGCTCAATTAATTTTGACACTTT 159
XX
RESULT 7
ABL07356/c
ID ABL07356 standard; cDNA; 4937 BP.
XX
XX ABL07356;
XX
DT 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 16550.
DE
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 16550.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
DR P-PSDB; ABB63253.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
XX Claim 1; SEQ ID NO 16550; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins (AAB57737-
CC AAB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at http://wipo.int/pub/publ/published_pct_sequences

SO Sequence 4937 BP; 1248 A; 1272 C; 1143 G; 1274 T; 0 U; 0 Other;

Query Match 51.5%; Score 24.2; DB 4; Length 4937;

Best Local Similarity 68.1%; Pred. No. 1.5e+02;

Matches 32; Conservative 1; Mismatches 14; Indels 0; Gaps 0;

QY 1 CCACGCGTTGATTAGAAATTAATTAATTGATCCACATTTCATTAATT 47
DB 2273 CAACAGTGGCTATGAGTATTAAGTTTAAACACACTTTCATTAATT 2227

RESULT 8

ABL32951/C
ID ABL32951 standard; DNA; 5173 BP.

AC ABL32951;

DT 26-MAR-2002 (first entry)

DE Human immune system associated gene SEQ ID NO: 924.

XX Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antiasthmatic; cytosinatic; nootropic;
KW neuroprotective; anti-HIV; anticoagulant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW ds.

XX Homo sapiens.

XX WO200200928-A2.

XX 03-JAN-2002.

PF 02-JUL-2001; 2001WO-EP007537.

XX 30-JUN-2000; 2000DE-01032529.

PR 01-SEP-2000; 2000DE-01043826.

XX (EPIG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI; 2002-130909/17.

PT Nucleic acid comprising fragment of chemically modified gene, useful for
PT diagnosis and treatment of diseases associated with abnormal cytosine
PT methylation.

PS Claim 1; SEQ ID NO 924; 32pp + Sequence Listing; German.

CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention

SO Sequence 5173 BP; 1274 A; 129 C; 1256 G; 2508 T; 0 U; 6 Other;

Query Match 51.5%; Score 24.2; DB 6; Length 5173;

Best Local Similarity 71.1%; Pred. No. 1.5e+02;
Matches 32; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 CCACGCGTTGATTAGAAATTAATTAATTGATCCACATTTCATTAATT 45
DB 3820 CCATATCCTTAATTAAAAATATTAATTTCAATTAATTCCTAATT 3776

RESULT 9

AAS36498
ID AAS36498 standard; DNA; 27681 BP.

AC AAS36498;

DT 17-DEC-2001 (first entry)

DE Human cardiovascular system antigen genomic DNA SEQ ID No 1998.

XX Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;
KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic; dog;
KW antirheumatic; antiproliferative; cytosinatic; cardiant; neuroprotective;
KW cerebroprotective; nootropic; antibacterial; vitruclide; fungicide; cancer;
KW ophthalmological; vulnery; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ds;
KW cerebrovascular disorder; nervous system disorder; bacterial infection;
KW fungal infection; viral infection; system disorder; endocrine disorder;
KW gastrointestinal disorder; renal disorder; respiratory disorder;
KW wound healing; skin aging; organ transplantation; tissue regeneration;
KW anti-infertility.

XX Homo sapiens.

XX WO200155321-A2.

XX 02-AUG-2001.

PF 17-JAN-2001; 2001WO-US001340.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220063P.

PR 16-AUG-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225477P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-022679P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226688P.

PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

RESULT 10
AAS36497
ID AAS36497 standard; DNA; 27681 BP.
XX
AC AAS36497;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human cardiovascular system antigen genomic DNA SEQ ID No 1997.
XX
KW Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;
KW chicken; sheep; immunosuppressive; anticharitic; vasotropic; dog;
KW antithematic; antiproliferative; cytostatic; cardiant; neuroprotective;
KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;
KW ophtalmological; vulnertary; gene therapy; autolimmune disease; neoplasm;
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ds;
KW cerebrovascular disorder; nervous system disorder; bacterial infection;
KW fungal infection; viral infection; ocular disorder; endocrine disorder;
KW gastrointestinal disorder; renal disorder; respiratory disorder;
KW wound healing; skin aging; organ transplantation; tissue regeneration;
KW anti-fertility.
XX
OS Homo sapiens.
XX
PN WO200155321-A2.
PD
XX 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001340.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-022679P.
PR 22-AUG-2000; 2000US-022681P.
PR 22-AUG-2000; 2000US-022688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231245P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231988P.
PR 14-SEP-2000; 2000US-023297P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234597P.
PR 25-SEP-2000; 2000US-0234598P.
PR 25-SEP-2000; 2000US-0234599P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.

PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251899P.
 PR 08-DEC-2000; 2000US-0251900P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-451930/48.
 XX
 PT New cardiovascular system related polynucleotides and polypeptides,
 PT useful for diagnosing, treating and/or preventing disorders of the
 PT cardiovascular system.
 XX
 PS Claim 1; SEQ ID NO 1997; 674bp; English.
 XX

CC Sequences AAS5741-AAS36942 represent genomic DNA molecules, which encode
 CC the cardiovascular system antigen polypeptides of the invention.
 CC Cardiovascular system antigens and their associated polynucleotides are
 CC useful in the diagnosis, treatment and prevention of various types of
 CC disorders in e.g. humans, mice, rabbits, goats, horses, cats, dogs,
 CC chickens or sheep. A pathological condition can be determined by
 CC detecting the presence or absence of a mutation in a cardiovascular
 CC system antigen polynucleotide. The treatable disorders include autoimmune
 CC diseases such as rheumatoid arthritis, hyperproliferative disorders such
 CC as neoplasms of the breast or liver, cardiovascular disorders such as
 CC cardiac arrest, cerebrovascular disorders such as cerebral ischaemia,
 CC nervous system disorders such as Alzheimer's disease, infections caused
 CC by bacteria, viruses and fungi, ocular disorders such as corneal
 CC infection, endocrine disorders such as premature labour and infertility,
 CC gastrointestinal disorders such as Crohn's disease, renal disorders such
 CC as glomerulonephritis and respiratory disorders such as asthma and
 CC pleurisy. The polypeptides can also be used to aid wound healing, to
 CC prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, to regenerate tissues and in chemotaxis. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX

Query Match 50.2%; Score 23.6; DB 4; Length 27681;
 Best Local Similarity 76.3%; Pred. No. 2.9e+02;
 Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 8 CTTGATTGAATTAATTAATTTGATCCATTTCCAATT 45
 Db 25518 CTTGATGAGAACTAATTGAGAACTATTTCATATT 25555

RESULT 11
 ID AK85843
 AC AK85843 standard; DNA; 27681 BP.
 XX

DT 07-NOV-2001 (first entry)
 XX
 DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:40655.
 XX
 KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytostatic; gene therapy; vaccine; metastasis; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200157182-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US001354.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 16-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226682P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0228287P.
 PR 01-SEP-2000; 2000US-0228343P.
 PR 01-SEP-2000; 2000US-0228344P.
 PR 01-SEP-2000; 2000US-0228345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0233398P.
 PR 14-SEP-2000; 2000US-0233399P.
 PR 14-SEP-2000; 2000US-023400P.
 PR 14-SEP-2000; 2000US-023401P.
 PR 14-SEP-2000; 2000US-0233063P.

PR	14-SEP-2000	2000US-02330645P
PR	14-SEP-2000	2000US-02330655P
PR	21-SEP-2000	2000US-02342233P
PR	21-SEP-2000	2000US-02342247P
PR	21-SEP-2000	2000US-02342977P
PR	25-SEP-2000	2000US-02349988P
PR	26-SEP-2000	2000US-02354844P
PR	27-SEP-2000	2000US-02358844P
PR	27-SEP-2000	2000US-02358856P
PR	29-SEP-2000	2000US-02363377P
PR	29-SEP-2000	2000US-02363567P
PR	29-SEP-2000	2000US-02363688P
PR	29-SEP-2000	2000US-02363699P
PR	29-SEP-2000	2000US-02363709P
PR	02-OCT-2000	2000US-02368028P
PR	02-OCT-2000	2000US-02370377P
PR	02-OCT-2000	2000US-02370388P
PR	02-OCT-2000	2000US-02370399P
PR	02-OCT-2000	2000US-02370409P
PR	13-OCT-2000	2000US-02393955P
PR	13-OCT-2000	2000US-02393977P
PR	20-OCT-2000	2000US-02409606P
PR	20-OCT-2000	2000US-02412211P
PR	20-OCT-2000	2000US-02417855P
PR	20-OCT-2000	2000US-02417866P
PR	20-OCT-2000	2000US-02417887P
PR	20-OCT-2000	2000US-02418088P
PR	20-OCT-2000	2000US-02418099P
PR	20-OCT-2000	2000US-02418266P
PR	01-NOV-2000	2000US-02464577P
PR	01-NOV-2000	2000US-02464747P
PR	08-NOV-2000	2000US-02464755P
PR	08-NOV-2000	2000US-02464776P
PR	08-NOV-2000	2000US-02464787P
PR	08-NOV-2000	2000US-02465233P
PR	08-NOV-2000	2000US-02465244P
PR	08-NOV-2000	2000US-02465255P
PR	08-NOV-2000	2000US-02465266P
PR	08-NOV-2000	2000US-02465277P
PR	08-NOV-2000	2000US-02465288P
PR	08-NOV-2000	2000US-02465522P
PR	08-NOV-2000	2000US-02466109P
PR	08-NOV-2000	2000US-02466110P
PR	08-NOV-2000	2000US-02466133P
PR	11-NOV-2000	2000US-02492077P
PR	11-NOV-2000	2000US-02492088P
PR	11-NOV-2000	2000US-02492099P
PR	11-NOV-2000	2000US-02492110P
PR	11-NOV-2000	2000US-02492111P
PR	11-NOV-2000	2000US-02492122P
PR	11-NOV-2000	2000US-02492133P
PR	11-NOV-2000	2000US-02492144P
PR	11-NOV-2000	2000US-02492155P
PR	11-NOV-2000	2000US-02492166P
PR	11-NOV-2000	2000US-02492177P
PR	11-NOV-2000	2000US-02492188P
PR	11-NOV-2000	2000US-02492244P
PR	11-NOV-2000	2000US-02492455P
PR	11-NOV-2000	2000US-02492466P
PR	11-NOV-2000	2000US-02503913P
PR	05-DEC-2000	2000US-02510309P
PR	05-DEC-2000	2000US-02519888P
PR	05-DEC-2000	2000US-02567199P
PR	06-DEC-2000	2000US-02514799P
PR	08-DEC-2000	2000US-02518566P
PR	08-DEC-2000	2000US-02518688P

PR	08-DEC-2000;	2000US-02518659P.	
PR	08-DEC-2000;	2000US-0251989P.	
PR	08-DEC-2000;	2000US-0251990P.	
PR	11-DEC-2000;	2000US-0254097P.	
PR	05-JAN-2001;	2001US-0259678P.	
XX			
PA	(HUMA-) HUMAN GENOME SCI INC.		
PL	Rosen CA, Barash SC, Ruben SM,		
DR	WPI; 2001-483426/52.		
XX			
PT	Nucleic acid encoding human immune/hematopoietic antigen polypeptides,		
PT	useful for preventing, diagnosing and/or treating cancers and metastasis.		
XX			
PS	Disclosure; SEQ ID NO 40655; 3071bp + Sequence Listing; English.		
XX			
CC	AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)		
CC	amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic		
CC	activity, and can be used in gene therapy and vaccine production. (I)		
CC	proteins and polynucleotides may be used in the prevention, diagnosis and		
CC	treatment of diseases associated with inappropriate (I) expression. For		
CC	example, they may be used to treat disorders associated with decreased		
CC	expression by rectifying mutations or deletions in a patient's genome		
CC	that affect the activity of (I) by expressing inactive proteins or to		
CC	supplement the patients own production of (I). Additionally, (I)		
CC	polynucleotides may be used to produce the secreted (I), by inserting the		
CC	nucleic acids into a host cell and culturing the cell to express the		
CC	protein. (I) proteins and polynucleotides may be used to prevent,		
CC	diagnose and treat immune/hematopoietic-related diseases, especially		
CC	cancers and cancer metastases of hematopoietic-derived cells. AAK44703		
CC	to AAK87694 represent human immune/hematopoietic antigen genomic		
CC	sequences from the present invention. AAK54942 to AAK54950 and AAK62169		
CC	represent sequences used in the exemplification of the present invention		
XX			
SO	Sequence 27681 BP; 7624 A; 6057 C; 5942 G; 8058 T; 0 U; 0 Other;		
Qy	Query Match	50.2%; Score 23.6; DB 4; Length 27681;	
	Best Local Similarity	76.3%; Pred. No. 2.9e+02;	
	Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;		
Db	25518 CTTGATGAGAGTAATTTTAGAACATTAATTTTCATATT 25555		
Oy	8 CTTGATTAAGAAATTAASITTCATCACCATTTCAAATT 45		
ID	AD647191 standard; DNA; 27681 BP.		
XX			
AC	AD647191;		
DT	29-JAN-2004 (first entry)		
XX			
DE	Human cardiovascular system related genomic DNA #757.		
XX			
KW	Human; cardiovascular system related polypeptide; cancer;		
KW	proliferative activity; fetal abnormality; developmental abnormality;		
KW	haematopoietic disorder; AIDS; autoimmune disease; rheumatoid arthritis;		
KW	inflammation; allergy; neurological disorder; Alzheimer's disease;		
KW	Parkinson's disease; cognitive disorder; schizophrenia; asthma;		
KW	skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;		
KW	cardiovascular disorder; angiotensin; kidney disorder;		
KW	gastrointestinal disorder; pregnancy-related disorder;		
KW	endocrine disorder; gene; ds.		
XX			
OS	Homo sapiens.		
XX			
FN	US2003059908-A1.		
XX			
PD	27-MAR-2003.		
XX			
PF	07-MAR-2002; 2002US-00091504.		

XX
PR 31-JAN-2000; 2000US-0179065P-
PR 04-FEB-2000; 2000US-0180628P-
PR 24-FEB-2000; 2000US-0184664P-
PR 02-MAR-2000; 2000US-0186550P-
PR 16-MAR-2000; 2000US-0189874P-
PR 17-MAR-2000; 2000US-0190076P-
PR 18-APR-2000; 2000US-0198123P-
PR 19-MAY-2000; 2000US-0205151P-
PR 07-JUN-2000; 2000US-0209467P-
PR 28-JUN-2000; 2000US-0214886P-
PR 30-JUN-2000; 2000US-0215135P-
PR 07-JUL-2000; 2000US-0216647P-
PR 11-JUL-2000; 2000US-0217487P-
PR 14-JUL-2000; 2000US-0217496P-
PR 26-JUL-2000; 2000US-0218290P-
PR 26-JUL-2000; 2000US-0220963P-
PR 14-AUG-2000; 2000US-0220964P-
PR 14-AUG-2000; 2000US-0224518P-
PR 14-AUG-2000; 2000US-0224519P-
PR 14-AUG-2000; 2000US-0225213P-
PR 14-AUG-2000; 2000US-0225214P-
PR 14-AUG-2000; 2000US-0225266P-
PR 14-AUG-2000; 2000US-0225267P-
PR 14-AUG-2000; 2000US-0225268P-
PR 14-AUG-2000; 2000US-0225270P-
PR 14-AUG-2000; 2000US-0225447P-
PR 14-AUG-2000; 2000US-0225757P-
PR 14-AUG-2000; 2000US-0225758P-
PR 14-AUG-2000; 2000US-0225759P-
PR 18-AUG-2000; 2000US-0226279P-
PR 22-AUG-2000; 2000US-0226681P-
PR 22-AUG-2000; 2000US-0226688P-
PR 22-AUG-2000; 2000US-0227182P-
PR 23-AUG-2000; 2000US-0227009P-
PR 30-AUG-2000; 2000US-0228924P-
PR 01-SEP-2000; 2000US-0229287P-
PR 01-SEP-2000; 2000US-0229343P-
PR 01-SEP-2000; 2000US-0229344P-
PR 01-SEP-2000; 2000US-0229345P-
PR 03-SEP-2000; 2000US-0229309P-
PR 05-SEP-2000; 2000US-0229513P-
PR 06-SEP-2000; 2000US-0230437P-
PR 06-SEP-2000; 2000US-0230438P-
PR 08-SEP-2000; 2000US-0231242P-
PR 08-SEP-2000; 2000US-0231243P-
PR 08-SEP-2000; 2000US-0231244P-
PR 08-SEP-2000; 2000US-0231413P-
PR 08-SEP-2000; 2000US-0231414P-
PR 08-SEP-2000; 2000US-0232080P-
PR 08-SEP-2000; 2000US-0232081P-
PR 12-SEP-2000; 2000US-0231968P-
PR 14-SEP-2000; 2000US-0232397P-
PR 14-SEP-2000; 2000US-0232398P-
PR 14-SEP-2000; 2000US-0232399P-
PR 14-SEP-2000; 2000US-0232400P-
PR 14-SEP-2000; 2000US-0232401P-
PR 14-SEP-2000; 2000US-0232403P-
PR 14-SEP-2000; 2000US-0232406P-
PR 14-SEP-2000; 2000US-0233064P-
PR 14-SEP-2000; 2000US-0233065P-
PR 21-SEP-2000; 2000US-0234223P-
PR 21-SEP-2000; 2000US-0234274P-
PR 25-SEP-2000; 2000US-0234997P-
PR 25-SEP-2000; 2000US-0234998P-
PR 26-SEP-2000; 2000US-0234984P-
PR 27-SEP-2000; 2000US-0235834P-
PR 27-SEP-2000; 2000US-0235835P-
PR 29-SEP-2000; 2000US-0236327P-
PR 29-SEP-2000; 2000US-0236367P-
PR 29-SEP-2000; 2000US-0236368P-
PR 29-SEP-2000; 2000US-0236369P-
PR 29-SEP-2000; 2000US-0236370P-
PT

PR 02-OCT-2000; 2000US-0236802P-
PR 02-OCT-2000; 2000US-0237037P-
PR 02-OCT-2000; 2000US-0237038P-
PR 02-OCT-2000; 2000US-0237039P-
PR 02-OCT-2000; 2000US-0237040P-
PR 13-OCT-2000; 2000US-0239935P-
PR 13-OCT-2000; 2000US-0239937P-
PR 20-OCT-2000; 2000US-0240960P-
PR 20-OCT-2000; 2000US-0241281P-
PR 20-OCT-2000; 2000US-0241785P-
PR 20-OCT-2000; 2000US-0241786P-
PR 20-OCT-2000; 2000US-0241787P-
PR 20-OCT-2000; 2000US-0241808P-
PR 20-OCT-2000; 2000US-0241809P-
PR 20-OCT-2000; 2000US-0241826P-
PR 01-NOV-2000; 2000US-0244617P-
PR 08-NOV-2000; 2000US-0244674P-
PR 08-NOV-2000; 2000US-0246475P-
PR 08-NOV-2000; 2000US-0246476P-
PR 08-NOV-2000; 2000US-0246477P-
PR 08-NOV-2000; 2000US-0246478P-
PR 08-NOV-2000; 2000US-0246523P-
PR 08-NOV-2000; 2000US-0246524P-
PR 08-NOV-2000; 2000US-0246525P-
PR 08-NOV-2000; 2000US-0246526P-
PR 08-NOV-2000; 2000US-0246527P-
PR 08-NOV-2000; 2000US-0246528P-
PR 08-NOV-2000; 2000US-0246533P-
PR 08-NOV-2000; 2000US-0246609P-
PR 08-NOV-2000; 2000US-0246610P-
PR 08-NOV-2000; 2000US-0246611P-
PR 08-NOV-2000; 2000US-0246613P-
PR 17-NOV-2000; 2000US-0249207P-
PR 17-NOV-2000; 2000US-0249208P-
PR 17-NOV-2000; 2000US-0249209P-
PR 17-NOV-2000; 2000US-0249210P-
PR 17-NOV-2000; 2000US-0249211P-
PR 17-NOV-2000; 2000US-0249212P-
PR 17-NOV-2000; 2000US-0249213P-
PR 17-NOV-2000; 2000US-0249214P-
PR 17-NOV-2000; 2000US-0249215P-
PR 17-NOV-2000; 2000US-0249216P-
PR 17-NOV-2000; 2000US-0249217P-
PR 17-NOV-2000; 2000US-0249218P-
PR 17-NOV-2000; 2000US-0249244P-
PR 17-NOV-2000; 2000US-0249245P-
PR 17-NOV-2000; 2000US-0249246P-
PR 17-NOV-2000; 2000US-0249264P-
PR 17-NOV-2000; 2000US-0249265P-
PR 17-NOV-2000; 2000US-0249297P-
PR 17-NOV-2000; 2000US-0249299P-
PR 17-NOV-2000; 2000US-0249300P-
PR 01-DEC-2000; 2000US-0250160P-
PR 01-DEC-2000; 2000US-0250391P-
PR 05-DEC-2000; 2000US-0250391P-
PR 05-DEC-2000; 2000US-0251030P-
PR 05-DEC-2000; 2000US-0256719P-
PR 05-DEC-2000; 2000US-0256719P-
PR 06-DEC-2000; 2000US-0251479P-
PR 08-DEC-2000; 2000US-0251865P-
PR 08-DEC-2000; 2000US-0251868P-
PR 08-DEC-2000; 2000US-0251869P-
PR 08-DEC-2000; 2000US-0251869P-
PR 08-DEC-2000; 2000US-0251989P-
PR 08-DEC-2000; 2000US-0251990P-
PR 11-DEC-2000; 2000US-0254097P-
PR 05-JAN-2001; 2001US-0259678P-
PR 17-JAN-2001; 2001US-00764869.
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Ruben SM, Barash SC;
XX WPI; 2003-743766/70.
XX
XX New cardiovascular system related polynucleotides and polypeptides,
PT

PT useful for preventing, treating, or ameliorating a medical condition.
PT such as cancer of cardiovascular tissues and cancer metastases.
XX
PS Claim 1, SEQ ID NO 1997, 262pp, English.
XX
XX The invention relates to human cardiovascular system related polypeptides
CC and the polynucleotides encoding them. The polypeptides, polynucleotides
CC and antibodies to the polypeptides are useful for diagnosing a
CC pathological condition or a susceptibility to a pathological condition,
CC for preventing, treating, or ameliorating a medical condition, such as
CC cancer of cardiovascular system tissues, proliferative disorders, foetal
CC and developmental abnormalities, haematopoietic disorders, diseases of
CC the immune system, AIDS, autoimmune diseases (e.g., rheumatoid
CC arthritis), inflammation, allergies, neurological disorders (e.g.,
CC Alzheimer's disease, Parkinson's disease), cognitive disorders,
CC schizophrenia, asthma, skin disorders (e.g., psoriasis), sepsis,
CC diabetes, atherosclerosis, cardiovascular disorders, angio-genic
CC disorders, kidney disorders, gastrointestinal disorders, pregnancy-
CC related disorders, endocrine disorders and infections. The nucleic acids
CC are also useful for chromosome identification, radiation hybrid mapping
CC or long-range restriction mapping. The polypeptides and polynucleotides
CC may also be used as food additives or preservatives to increase or
CC decrease storage capabilities, fat content or other nutritional
CC components. This sequence represents human cardiovascular system related
CC genomic DNA of the invention.
XX
SQ Sequence 27681 BP, 7624 A, 6057 C, 5942 G, 8058 T, 0 U, 0 Other;
Query Match 50.2%; Score 23.6; DB 10; Length 27681;
Best Local Similarity 76.3%; Pred. No. 2.9e+02;
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 8 CTTGATTAGAAATTAASSTTGATCACCATTTCAAATT 45
DB 25518 CTTGATGAGAGATTAATTTAGAACATTATTTCATATT 25555
RESULT 13
ADE47192
ID ADE47192 standard; DNA; 27681 BP.
XX
AC ADE47192;
XX
XX 29-JAN-2004 (first entry)
XX
XX Human cardiovascular system related genomic DNA #758.
DE
XX
XX Human: cardiovascular system related polypeptide; cancer;
KW proliferative disorder; foetal abnormality; developmental abnormality;
KW haematopoietic disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW cardiovascular disorder; angio-genic disorder; kidney disorder;
KW gastrointestinal disorder; pregnancy-related disorder;
KW endocrine disorder; gene; ds.
XX
XX Homo sapiens.
OS
XX
PN US2003059908-A1.
PD
XX
XX 27-MAR-2003.
XX
XX 07-MAR-2002; 2002US-00091504.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-02331968P.
PR 14-SEP-2000; 2000US-0233297P.
PR 14-SEP-2000; 2000US-0233298P.
PR 14-SEP-2000; 2000US-0233299P.
PR 14-SEP-2000; 2000US-0233400P.
PR 14-SEP-2000; 2000US-0233401P.
PR 14-SEP-2000; 2000US-0233403P.
PR 14-SEP-2000; 2000US-0233404P.
PR 14-SEP-2000; 2000US-0233405P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234497P.
PR 25-SEP-2000; 2000US-0234498P.
PR 26-SEP-2000; 2000US-0234484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.

PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235634P.
PR 27-SEP-2000; 2000US-0235636P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241211P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0251719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764869.
PR 07-MAR-2002; 2002US-00091504.
(HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2004-081713/08.
DR
XX
XX
PT New cardiovascular system-related nucleic acid molecule, useful for
PT diagnosing, preventing or treating diseases of the cardiovascular system,
PT and in chromosome mapping, drug screening or in pharmacogenomics.
XX
XX
PS Disclosure: SEQ ID NO 1998; 262dp; English.
XX
XX The invention relates to an isolated nucleic acid molecule encoding a
CC human cardiovascular system associated polypeptide (or antigens), or its
CC fragment. Also included recombinant vectors, recombinant host cells, an
CC isolated human cardiovascular system associated polypeptide (including
CC its fragment, allelic variant, species homologue or epitope), an isolated
CC antibody that binds specifically to a human cardiovascular system
CC associated polypeptide, diagnosing a pathological condition or
CC susceptibility to a pathological condition (comprising determining the
CC presence or absence of a mutation in human cardiovascular system
CC associated nucleic acid and diagnosing a condition based on the presence
CC or absence of the mutation), identifying a binding partner to human
CC cardiovascular system associated polypeptides, the gene corresponding to

CC the human cardiovascular system associated cDNA sequence and identifying
CC an activity in a biological assay comprising expressing the human
CC cardiovascular system associated cDNA in a cell, isolating the
CC supernatant, detecting an activity in a biological assay and identifying
CC the protein in the supernatant having the activity. The human
CC cardiovascular system associated nucleic acids and polypeptides are used
CC to prevent, treat or ameliorate a medical condition (for example in
CC humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep), for
CC example autoimmune diseases such as rheumatoid arthritis,
CC hyperproliferative disorders, for example neoplasms of the breast or
CC liver, cardiovascular disorders, for example cardiac arrest,

Query Match 50.2%; Score 23.6; DB 13; Length 27681;

Best Local Similarity 76.3%; Pred. No. 2.9e+02;

Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 8 CTTGATTGAATTAATTTGATCACCATTTCGAATT 45
Db 25518 CTTGATGAGAAAGTAATTTAGAACATTAATTCATATT 25555

RESULT 15

ADJ08609 standard; DNA; 27681 BP.

ADJ08609;

04-NOV-2004 (first entry)

Human cardiovascular system associated polypeptide-related DNA SeqID1997.

KM autoimmune disease; rheumatoid arthritis; hyperproliferative disorder;
KM breast neoplasms; liver neoplasm; cardiovascular disorder;
KM cardiac arrest; cerebrovascular disorder; cerebral ischaemia;
KM angiogenesis; nervous system disorder; Alzheimer's disease; infection;
KM ocular disorder; corneal infection; skin aging; sunburn;
KM epithelial cell proliferation; cell culture; tissue regeneration; chemotaxis;
KM organ transplantation; cell culture; tissue regeneration; chemotaxis;
KM food additive; preservative; cardiovascular system associated antigen;
KM nuclear factor kappaB; NFkappaB; promoter element; human; ds.

XX Homo sapiens.

XX US2004005575-A1.

PD 08-JAN-2004.

PF 26-AUG-2002; 2002US-00227577.

XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225265P.
PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 18-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227099P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0228287P.
PR 01-SEP-2000; 2000US-0228343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232081P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0233977P.
PR 14-SEP-2000; 2000US-0233989P.
PR 14-SEP-2000; 2000US-0233999P.
PR 14-SEP-2000; 2000US-0234000P.
PR 14-SEP-2000; 2000US-0234011P.
PR 14-SEP-2000; 2000US-0234063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234988P.
PR 26-SEP-2000; 2000US-0234984P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239353P.
PR 13-OCT-2000; 2000US-0239373P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0244674P.
PR 08-NOV-2000; 2000US-0244675P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.

PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0250100P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764869.
PR 07-MAR-2002; 2002US-00091504.

(HUMA-) HUMAN GENOME SCI INC.
XX
PA
PI Rosen CA, Ruben SM, Barash SC;
DR WPI; 2004-081713/08.
XX
XX
PT New cardiovascular system-related nucleic acid molecule, useful for
PT diagnosing, preventing or treating diseases of the cardiovascular system,
XX and in chromosome mapping, drug screening or in pharmacogenomics.
XX
PS Disclosure; SEQ ID NO 1997; 262bp; English.
XX
XX
CC The invention relates to an isolated nucleic acid molecule encoding a
CC human cardiovascular system associated polypeptide (or antigens), or its
CC fragment. Also included recombinant vectors, recombinant host cells, an
CC isolated human cardiovascular system associated polypeptide (including
CC its fragment, allelic variant, species homologue or epitope), an isolated
CC antibody that binds specifically to a human cardiovascular system
CC associated polypeptide, diagnosing a pathological condition or
CC susceptibility to a pathological condition (comprising determining the
CC presence or absence of a mutation in human cardiovascular system
CC associated nucleic acid and diagnosing a condition based on the presence
CC or absence of the mutation), identifying a binding partner to human
CC cardiovascular system associated polypeptides, the gene corresponding to
CC the human cardiovascular system associated cDNA sequence and identifying
CC an activity in a biological assay comprising expressing the human
CC cardiovascular system associated cDNA in a cell, isolating the
CC supernatant, detecting an activity in a biological assay and identifying
CC the protein in the supernatant having the activity. The human
CC cardiovascular system associated nucleic acids and polypeptides are used
CC to prevent, treat or ameliorate a medical condition (for example in

CC humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep), for
CC example autoimmune diseases such as rheumatoid arthritis,
CC hyperproliferative disorders, for example neoplasms of the breast or
CC liver, cardiovascular disorders, for example cardiac arrest.

Query Match 50.2%; Score 23.6; DB 13; Length 27681;
Best Local Similarity 76.3%; Pred. No. 2.9e+02;
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 8 CTTGATTAGAAATTAATTTGATCACCATTTCGAATT 45
DB 25518 CTTGATGAGAGTAATTTAGAACATTATTTTCATATT 25555

RESULT 16
AAH93984/C
ID AAH93984 standard; cDNA; 342 BP.
XX
XX
AC AAH93984;
XX
DT 05-OCT-2001 (first entry)
XX
XX
DE Human foetal cDNA, SEQ ID NO: 513.
XX
XX Human; foetal protein; cytosolic; immunosuppressive; immunostimulant;
XX neurotropic; neuroprotective; thrombolytic; osteopathic; antiinflammatory;
XX gene therapy; antisense therapy; cancer; immune disorder;
XX growth disorder; osteoporosis; thrombolytic disorder;
XX nervous system disorder; inflammation; expressed sequence tag; EST; ss.
XX
OS Homo sapiens.
XX
XX MO200155339-A2.
PN
XX
PD 02-AUG-2001.
XX
XX 25-JAN-2001; 2001MO-US002723.
PF
XX
XX 25-JAN-2000; 2000US-00491404.
PR 15-SEP-2000; 2000US-00663870.
PR 06-NOV-2000; 2000US-00707351.
XX
XX
PA (HYSE-) HYSEQ INC.
XX
XX
PI Yeung G, Ford JE, Boyle BJ, Arterburn MC, Dormanac RA, Tang YT;
PI Liu C, Asundi V, Zhou P, Wehrman T;
XX
XX
DR WPI; 2001-465571/50.
XX
XX P-PSDB; AAM06309.
XX
XX
PT Novel fetal proteins useful for the treatment and diagnosis of diseases
PT associated with dysfunction of the protein e.g. cancers, immune
PT disorders, growth disorders, thrombolytic disorders, nervous system
PT disorders and inflammation.
XX
XX
PS Claim 1; Page 365; 715bp; English.
XX
XX
CC The invention relates to novel foetal polypeptides encoded by
CC polynucleotides comprising one of 477 sequences fully defined in the
CC specification. The foetal polynucleotides and polypeptides are useful in
CC the treatment and diagnosis of diseases such as cancers, immune
CC disorders, growth disorders (e.g. osteoporosis), thrombolytic disorders,
CC nervous system disorders and inflammation. The present sequence was
CC assembled using an expressed sequence tag (EST) found to be expressed in
XX human foetal tissue cDNA libraries as the seed
XX
XX
SQ Sequence 342 BP; 125 A; 46 C; 57 G; 114 T; 0 U; 0 Other;

Query Match 49.8%; Score 23.4; DB 5; Length 342;
Best Local Similarity 69.8%; Pred. No. 2.2e+02;
Matches 30; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 5 AGGCTTGATTAGAAATTAATTTGATCACCATTTCGAATTTT 47

Db 99 AGGATTGATTCYAACTTAACATTAACTACGACATTATTAATATTTT 57
||| ||||| ||: || |||| ||| |||||
RESULT 17
AAK70686/c
ID AAK70686 standard; DNA; 34917 BP.
XX
XX AAK70686;
AC
XX
XX 06-NOV-2001 (first entry)
DT
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:25498.
DE
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW Cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
OS
XX WO200157182-A2.
PN
PD 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001354.
PF
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 18-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 30-AUG-2000; 2000US-0227094P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239355P.
PR 13-OCT-2000; 2000US-023937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0244674P.
PR 08-NOV-2000; 2000US-0244675P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246522P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.

PR	17-NOV-2000;	2000US-0249299P.
PR	17-NOV-2000;	2000US-0249300P.
PR	01-DEC-2000;	2000US-0250160P.
PR	01-DEC-2000;	2000US-0250391P.
PR	05-DEC-2000;	2000US-0251030P.
PR	05-DEC-2000;	2000US-0251988P.
PR	05-DEC-2000;	2000US-0256719P.
PR	06-DEC-2000;	2000US-0251479P.
PR	08-DEC-2000;	2000US-0251866P.
PR	08-DEC-2000;	2000US-0251868P.
PR	08-DEC-2000;	2000US-0251989P.
PR	08-DEC-2000;	2000US-0251990P.
PR	11-DEC-2000;	2000US-0254097P.
PR	05-JAN-2001;	2001US-0259678P.
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
PI	Rosen CA, Barash SC, Ruben SM;	
DR	WPI; 2001-483426/52.	
XX		
PT	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,	
PT	useful for preventing, diagnosing and/or treating cancers and metastasis.	
XX		
PS	Disclosure; SEQ ID NO 25498; 3071bp + Sequence Listing; English.	
XX		
CC	AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)	
CC	amino acid sequences given in AAM82170 to AAM91921. (I) have cytosolic	
CC	activity, and can be used in gene therapy and vaccine production. (I)	
CC	proteins and polynucleotides may be used in the prevention, diagnosis and	
CC	treatment of diseases associated with inappropriate (I) expression. For	
CC	expression, they may be used to treat disorders associated with decreased	
CC	expression by rectifying mutations or deletions in a patient's genome	
CC	that affect the activity of (I) by expressing inactive proteins or to	
CC	supplement the patients own production of (I). Additionally, (I)	
CC	polynucleotides may be used to produce the secreted (I), by inserting the	
CC	nucleic acids into a host cell and culturing the cell to express the	
CC	protein. (I) proteins and polynucleotides may be used to prevent,	
CC	diagnose and treat immune/haematopoietic-related diseases, especially	
CC	cancers and cancer metastases of haematopoietic-derived cells. AAK64703	
CC	to AAK87654 represent human immune/haematopoietic antigen genomic	
CC	sequences from the present invention. AAK54942 to AAK54950 and AAM82169	
CC	represent sequences used in the exemplification of the present invention	
CC		
XX		
SQ	Sequence 34917 BP; 10788 A; 6622 C; 7157 G; 10350 T; 0 U; 0 Other;	
OY	Query Match	49.8%; Score 23.4; DB 4; Length 34917;
	Best Local Similarity	69.8%; Pred. No. 3.5e+02;
D8	Matches 30; Conservative 1; Mismatches 12; Indels 0; Gaps 0;	
OY	5 AGGCTGATTAGAAATTAATASTTTGTTCACCACTTTTCAATTTT 47	
D8	15031 AGCCTTCATTAATTATTAACAATTGTGGAAATTAATTAAATTAT 14989	
RESULT 18		
ACF69221/c		
ID	ACF69221 standard; DNA; 564 BP.	
XX		
AC	ACF69221;	
XX		
D7	20-NOV-2003 (first entry)	
DE	Photorhabdus luminescens nucleotide sequence #7688.	
XX		
KM	Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;	
KM	detection; food; gene expression; plant; animal; microorganism; toxin;	
KM	antibiotic; bioplasticide; virulence factor; disease model; plague;	
XX	whooping cough; gene; de.	
XX		
OS	Photorhabdus luminescens.	
XX		

PN	MO200294867-A2.
XX	
PD	28-NOV-2002.
XX	
PF	07-FEB-2002; 2002MO-IB003040.
XX	
PR	07-FEB-2001; 2001FR-00001659.
XX	
PA	(INSP) INST PASTEUR.
XX	
PA	(CNRS) CNRS CENT NAT RECH SCI.
XX	
PI	Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
XX	
PI	Buchrieser C;
XX	
DR	WPI; 2003-148459/14.
XX	
PT	Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
XX	
PT	useful e.g. as therapeutic antimicrobials and agricultural pesticides.
XX	
PS	Claim 2; SEQ ID NO 7688; 1205BP; French.
XX	
CC	The invention relates to the isolation of genes and their encoded
XX	
CC	proteins from Photorhabdus luminescens. The isolated sequences are
XX	
CC	sources of probes and primers for detecting the genome of P. luminescens
XX	
CC	and related species; to study polymorphisms; for gene analysis and for
XX	
CC	detection/amplification of the genes. Antibodies (Ab) raised against the
XX	
CC	polypeptides encoded by the genes are used for detection/identification
XX	
CC	of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
XX	
CC	carry a gene-containing vector are used to select compounds that
XX	
CC	modulate, regulate, induce or inhibit expression of the genes in plants,
XX	
CC	animals or microorganisms other than P. luminescens and are able to alter
XX	
CC	response or sensitivity to toxins and antibiotics produced by P.
XX	
CC	luminescens. Cells transformed to express the genes are useful for
XX	
CC	recombinant production of the proteins, particularly toxins and
XX	
CC	antibacterials useful as insecticides, bactericides and fungicides. The
XX	
CC	therapeutically (to treat microbial infection by bacteria or fungi that
XX	
CC	are sensitive to P. luminescens-encoded toxins or antibiotics) and as
XX	
CC	biopesticides. Other uses of the genes and the proteins are as virulence
XX	
CC	factors and for identifying targets of human diseases for which P.
XX	
CC	luminescens is a model (particularly plague and whooping cough). This
XX	
CC	sequence represents one of the isolated P. luminescens genes
XX	
SQ	Sequence 564 BP; 182 A; 127 C; 111 G; 144 T; 0 U; 0 Other;
XX	
Query Match	49.4%; Score 23.2; DB 10; Length 564;
XX	
Best Local Similarity	73.7%; Pred. No. 2.7e+02;
XX	
Matches	28; Conservative 1; Mismatches 9; Indels 0; Gaps 0.
XX	
QY	3 ACAGGCTTGATTAGAAATTAASITTTGATTCACCATTTTC 40
XX	
DB	215 ACAGGCTTGTTAGATTGACACTTTCATCCGACATCTTC 178
XX	
RESULT 19	
XX	ABN67694
ID	ABN67694 standard; DNA; 939 BP.
XX	
AC	ABN67694;
XX	
DT	01-JUL-2002 (first entry)
XX	
DE	Streptococcus polynucleotide SEQ ID NO 3301.
XX	
XX	Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW	group A streptococcus; Streptococcus pyogenes; antibacterial; gene;
XX	ant inflammatory; infection; vaccine; meningitis; gene therapy; ds.
OS	Streptococcus pyogenes.
XX	
FN	MO200234771-A2.
XX	
PD	02-MAY-2002.

XX 29-OCT-2001; 2001WO-GB004789.
XX
XX 27-OCT-2000; 2000GB-00026333.
PR 24-NOV-2000; 2000GB-00028727.
PR 07-MAR-2001; 2001GB-00005640.
XX
XX (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX
PI Telford J, Maignani V, Margaric Y Rost, Grandi G, Fraser C;
PI Tettein H;
XX WPI; 2002-352536/38.
DR P-PSDB; ABP27063.
XX
XX New Streptococcus protein for the treatment or prevention of infection or
PT disease caused by Streptococcus bacteria, such as meningitis, and for
PT detecting a compound that binds to the protein.
PS Claim 7; Page 3484; 45255P; English.
XX
XX The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and anti-inflammatory
CC activity. (1), nucleic acids encoding (1), ABN6044-ABN71526 and
CC antibodies that bind (1) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (1) are used to detect Streptococcus in a
CC biological sample. (1) is used to determine whether a compound binds to
CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (1) may be used to recombinantly produce (1) and may be
CC used in gene therapy. Antibodies to (1) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins
XX
SQ Sequence 939 BP; 257 A; 172 C; 201 G; 309 T; 0 U; 0 Other;
Query Match 49.4%; Score 23.2; DB 6; Length 939;
Best Local Similarity 70.5%; Pred. No. 2.8e+02;
Matches 31; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 3 ACAGGCTGATTAAGAAATTAATTGATCACCATTTCGAATTT 46
DB 791 ACAAGCATGACAAAAGCAGTTTGATTACCATTTTGACATGT 834
RESULT 20
ACD13402
ID ACD13402 standard; cDNA; 1978 BP.
XX
XX ACD13402;
DT 13-AUG-2003 (first entry)
XX
XX Human DNA encoding a p53 modifier, SEQ ID 74.
XX
XX Homo sapiens.
XX OS
XX WO200299122-A1.
XX
XX 12-DEC-2002.
PD 03-JUN-2002; 2002WO-US017382.
XX
XX

XX 05-JUN-2001; 2001US-0296076P.
PR 10-OCT-2001; 2001US-0328605P.
PR 15-FEB-2002; 2002US-0357253P.
XX
XX (EXEL-) EXELIXIS INC.
XX
XX Friedman I, Plozman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
PI WPI; 2003-156859/15.
DR P-PSDB; ABO07229.
XX
XX Identifying modulators of the p53 pathway for use in treating apoptotic
PT or cell proliferation disorders, comprises screening for agents that
PT modulate activity of a human ortholog of genes that modify the p53
PT pathway in Drosophila.
XX
PS Example 2; Page 357-358; 678pp; English.
XX
XX The invention relates to identifying (M1) a candidate p53 pathway
CC modulating agent, by contacting an assay system comprising a purified HM
CC polypeptide (human orthologue of genes that modify the p53 pathway in
CC Drosophila) or nucleic acid with a test agent under conditions, where but
CC for the presence of the test agent, the system provides a reference
CC activity, and detecting a test agent-biased activity of the assay system.
CC Also included are modulating (M2) a p53 pathway of a cell (comprising
CC contacting a cell defective in p53 function with a candidate modulator
CC that specifically binds to a HM polypeptide comprising an HM amino acid
CC sequence, where p53 function is restored), modulating (M3) a p53 pathway
CC in a mammalian cell (comprising contacting the cell with an agent that
CC specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4)
CC a disease in a patient (comprising: (a) obtaining a biological sample
CC from the patient; (b) contacting the sample with a probe for HM
CC expression; (c) comparing the results with a control; and (d) determining
CC whether the comparison indicates a likelihood disease). (M1) is useful
CC for identifying modulators of the p53 pathway. A probe for HM expression
CC is useful for diagnosing breast, colon, kidney, lung and ovarian cancer.
CC in a patient, where the cancer has greater than 25 % expression level.
CC Modulators identified by (M1) are useful in a variety of diagnostic and
CC therapeutic applications, where disease or disorder prognosis is related
CC to defects in the p53 pathway, such as, angiogenesis' apoptotic or cell
CC proliferation disorders (e.g. cancer). Another two new methods (M2 and
CC M3) are useful for modulating the p53 pathway of a cell, thus restoring
CC the p53 function of the cell, so that the cell undergoes normal
CC proliferation or progression through the cell cycle. (M2) and (M3) are
CC also useful for treating defects in the p53 pathway such as angiogenic,
CC apoptotic or cell proliferation disorders. The present sequence is an HM
CC nucleic acid encoding a p53 pathway modifying protein
XX
SQ Sequence 1978 BP; 527 A; 476 C; 556 G; 419 T; 0 U; 0 Other;
Query Match 49.4%; Score 23.2; DB 8; Length 1978;
Best Local Similarity 83.3%; Pred. No. 3e+02;
Matches 25; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 10 TGATTAGAAATTAAGTTGATCACCATTTT 39
DB 1939 TGACTTGAATTAATTAATTGACACAAATTTT 1968
RESULT 21
ACN41054
ID ACN41054 standard; cDNA; 1978 BP.
XX
XX ACN41054;
AC
XX 18-NOV-2004 (first entry)
DT
XX
XX Tumour-associated antigenic target (TAT) cDNA DNA270496, SEQ ID NO:6239.
DE
XX Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KM tumour; diagnosis; cell proliferative disorder; breast cancer;
KM colorectal cancer; lung cancer; ovarian cancer; liver cancer;
XX

KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic; gene; ss.
XX Homo sapiens.
XX MO2004030615-A2.
XX 15-APR-2004.
XX 29-SEP-2003; 2003WO-US028547.
XX 02-OCT-2002; 2002US-0414971P.
XX (GETH) GENENTECH INC.
XX Wu TD, Zhang Z, Zhou Y;
XX WPI: 2004-347921/32.
XX P-PSDB; ABM82429.
XX New tumor-associated antigenic target polypeptides and nucleic acids,
XX useful in preparing a medicament for treating or detecting a
XX proliferative disorder, e.g. breast, lung, colorectal, ovarian or
XX prostate cancer or tumor.
XX Claim 1; SEQ ID NO 6239; 7273bp; English.
XX
XX The invention relates to human tumour-associated antigenic target (TAT)
XX polypeptides, and their related nucleic acids. The TAT polypeptides are
XX overexpressed in cancer tissues compared to normal tissues, and may thus
XX serve as effective targets for the diagnosis and treatment of cancer in
XX mammals. The invention also relates to nucleic acid and polypeptide
XX sequences at least 80% identical to the TAT nucleic acids and
XX polypeptides; expression vectors and host cells comprising a TAT nucleic
XX acid; an antibody specific for a TAT polypeptide; a peptide or organic
XX molecule which binds to a TAT polypeptide; fusion proteins comprising a
XX TAT polypeptide; and methods and compositions for the treatment or
XX diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
XX antibodies, antagonists, binding molecules and compositions are useful
XX for diagnosing or treating a cell proliferative disorder associated with
XX increased TAT expression, particularly cancers such as breast cancer,
XX colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
XX cancer, pancreatic cancer, cervical cancer, cancers of the central
XX nervous system, melanoma and leukaemia. TAT nucleic acids may further be
XX used as hybridisation probes, in chromosome and gene mapping, in
XX chromosome identification and in gene therapy. The present sequence
XX represents a TAT nucleic acid of the invention
XX
XX Sequence 1978 BP; 527 A; 476 C; 556 G; 419 T; 0 U; 0 Other;
SQ
XX
XX Query Match 49.4%; Score 23.2; DB 13; Length 1978;
XX Best Local Similarity 83.3%; Pred. No. 3e+02;
XX Matches 25; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 10 TGATTGAAATTAAGTTGATCACCATTTT 39
DB 1939 TGACTTGAATTAACCTTGACACCAATTTT 1968
RESULT 22
ADJ40809
XX ADJ40809 standard; cDNA; 2000 BP.
XX
XX ADJ40809;
AC
XX 06-MAY-2004 (first entry)
XX
XX Plant CDNA #1809.
DE
XX
XX Plant; gene; ss; transcription; plant genome augmentation; cereal;
KW soybean; alfalfa; sunflower; canola; cotton; peanut; tobacco; sugar beet;

KW maize; barley; sorghum; rice; wheat; crop plant; insecticide resistance;
KW stress tolerance; salt tolerance; cold tolerance; drought tolerance;
KW plant nutrition; apical dominance; dwarfism; early flowering; antiviral;
KW antifungal.
XX Eukaryota.
XX
XX US2004016025-A1.
XX 22-JAN-2004.
XX 26-SEP-2002; 2002US-00260238.
XX
XX 26-SEP-2001; 2001US-0325277P.
XX 26-SEP-2001; 2001US-0325448P.
XX 04-APR-2002; 2002US-0370620P.
XX
XX (BUDW/) BUDWORTH P.
XX (MOUG/) MOUGHAMER T.
XX (BRIG/) BRIGGS S P.
XX (COOP/) COOPER B.
XX (GLAZ/) GLAZEBROOK J.
XX (GOFF/) GOFF S A.
XX (KATG/) KATAGIRI F.
XX (KREP/) KREPS J.
XX (PROV/) PROVART N.
XX (RICK/) RICHE D.
XX (ZHU/) ZHU T.
XX
XX Budworth P, Moughamer T, Briggs SP, Cooper B, Glazebrook J;
XX Goff SA, Katagiri F, Krep J, Provart N, Riche D, Zhu T;
XX WPI: 2004-190374/18.
XX
XX New rice promoter, useful for manipulating crop plants to alter or
XX improve phenotypic characteristics, e.g. produce large quantities of oil
XX or proteins, resistance to insecticides, virus or fungi, stress tolerance
XX or high nutritional value.
XX
XX Claim 26; SEQ ID NO 1809; 230bp; English.
XX
XX The invention relates to plant nucleotide sequences that direct seed-,
XX leaf- and/or stem-, panicle-, root- or pollen-specific or -preferential
XX or constitutive transcription of an operatively linked nucleic acid
XX segment. The invention also relates to a method for augmenting a plant
XX genome and a method of identifying a gene, where its expression is
XX altered in the seed, leaf, stem, panicle, pollen, root or is constitutive
XX in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,
XX canola, cotton, peanut, tobacco or sugar beet, preferably maize, barley,
XX sorghum, rice or wheat. The polynucleotides and the polypeptides they
XX encode are useful for manipulating crop plants to alter or improve
XX phenotypic characteristics, to produce large quantities of oil or
XX proteins, to incur resistance to insecticides, viruses or fungi, and to
XX incur stress tolerance (e.g. salt, cold or drought) to ensure the plants
XX have a high nutritional value with reduced apical dominance or dwarfism,
XX early flowering or altered metabolic pathways. This sequence represents a
XX plant nucleic acid of the invention. Note: The sequence data for this
XX patent did not form part of the printed specification but was obtained in
XX electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 2000 BP; 614 A; 364 C; 405 G; 617 T; 0 U; 0 Other;
SQ
XX
XX Query Match 49.4%; Score 23.2; DB 12; Length 2000;
XX Best Local Similarity 77.8%; Pred. No. 3e+02;
XX Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 12 ATTGAAATTAAGTTGATCACCATTTTCAATTTT 47
DB 788 ACTTAATATCACATTTAAACACCATTTCATAATTTT 823
RESULT 23
ADM02167

ID ADM02167 standard; CDNA; 2347 BP.
XX
XX ADM02167;
AC
XX 20-MAY-2004 (first entry)
DT
XX Human CDNA of the invention SEQ ID NO:852.
DE
XX ss; gene; human; gene therapy; diagnostic marker; pharmaceutical.
XX
XX Homo sapiens.
OS
XX EPI347046-A1.
FN
XX 24-SEP-2003.
XX
XX 12-APR-2002; 2002EP-00008400.
PF
XX 22-MAR-2002; 2002JP-00137785.
PR
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
XX Isogai T, Sugiyama T, Otsuki T, Makamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y;
XX MPI; 2003-723558/69.
DR P-PSDB; ADM04610.
XX
XX New polynucleotides and polypeptides are useful in gene therapy, for
PT developing a diagnostic marker or medicines for regulating their
PT expression and activity, or as a target of gene therapy.
XX
XX Claim 1; SEQ ID NO 852; 305pp; English.
PS
XX The invention relates to a novel human polynucleotide and the encoded
CC polypeptide. A polynucleotide of the invention ADM06202-ADM06773 is useful
CC as a primer for synthesizing the polynucleotide or as a probe for
CC detecting the polynucleotide. The polynucleotide ADM01316-ADM03758 are
CC useful in gene therapy, for developing a diagnostic marker or medicines
CC for regulating their expression and activity, or as a target of gene
CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
CC are useful as pharmaceutical agents. The present sequence represents a
CC CDNA sequence of the invention.
XX
XX Sequence 2347 BP; 583 A; 516 C; 580 G; 668 T; 0 U; 0 Other;
SQ
Query Match 49.4%; Score 23.2; DB 11; Length 2347;
Best Local Similarity 83.3%; Pred. No. 3.1e+02;
Matches 25; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 10 TGATTGAAATTAATTTGATCACCATTTT 39
DB 2308 TGACTGAAATTAACCTTGAAACACATTTT 2337

RESULT 24
ABL28314
ID ABL28314 standard; DNA; 3212 BP.
XX
XX ABL28314;
AC
XX 26-MAR-2002 (first entry)
DT
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 36415.
DE
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; de.
XX
XX Drosophila melanogaster.
OS
XX
XX WO200171042-A2.
FN

XX 27-SEP-2001.
PD
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
PF
XX 11-JUL-2000; 2000US-00614150.
PR
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI MPI; 2001-656860/75.
DR
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Claim 1; SEQ ID NO 36415; 21pp + Sequence Listing; English.
PS
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 3212 BP; 849 A; 631 C; 681 G; 1051 T; 0 U; 0 Other;
SQ
Query Match 49.4%; Score 23.2; DB 4; Length 3212;
Best Local Similarity 70.5%; Pred. No. 3.2e+02;
Matches 31; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 4 CAGGCTTGATTGAAATTAATTTGATCACCATTTCATTTT 47
DB 2108 CACATTGATTGAAATTAATGATTTTGTGCTACCTAATTATTTT 2151

RESULT 25
AAK89062/C
ID AAK89062 standard; DNA; 5468 BP.
XX
XX AAK89062;
AC
XX 05-NOV-2001 (first entry)
DT
XX Human digestive system antigen genomic sequence SEQ ID NO: 2638.
DE
XX Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW digestive system disorder; Meckel's diverticulum; ds.
XX
XX Homo sapiens.
OS
XX
XX WO200155314-A2.
FN
XX 02-AUG-2001.
PD
XX 17-JAN-2001; 2001WO-US001324.
PF
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR

PR	8-JUN-2000	2000US-0214986P
PR	30-JUL-2000	2000US-0215135P
PR	07-JUL-2000	2000US-0216647P
PR	07-JUL-2000	2000US-0216880P
PR	11-JUL-2000	2000US-0217487P
PR	11-JUL-2000	2000US-0217496P
PR	14-JUL-2000	2000US-0218290P
PR	26-JUL-2000	2000US-0220963P
PR	26-JUL-2000	2000US-0220964P
PR	14-AUG-2000	2000US-0224518P
PR	14-AUG-2000	2000US-0225268P
PR	14-AUG-2000	2000US-0225270P
PR	14-AUG-2000	2000US-0225447P
PR	14-AUG-2000	2000US-0225575P
PR	14-AUG-2000	2000US-0225758P
PR	14-AUG-2000	2000US-0225759P
PR	18-AUG-2000	2000US-0226279P
PR	22-AUG-2000	2000US-0226661P
PR	22-AUG-2000	2000US-0226668P
PR	22-AUG-2000	2000US-0227182P
PR	30-AUG-2000	2000US-0227700P
PR	30-AUG-2000	2000US-0228924P
PR	01-SEP-2000	2000US-0229287P
PR	01-SEP-2000	2000US-0229343P
PR	01-SEP-2000	2000US-0229344P
PR	01-SEP-2000	2000US-0229345P
PR	05-SEP-2000	2000US-0229509P
PR	05-SEP-2000	2000US-0229513P
PR	06-SEP-2000	2000US-0230437P
PR	06-SEP-2000	2000US-0230438P
PR	08-SEP-2000	2000US-0231242P
PR	08-SEP-2000	2000US-0231243P
PR	08-SEP-2000	2000US-0231244P
PR	08-SEP-2000	2000US-0231413P
PR	08-SEP-2000	2000US-0231414P
PR	08-SEP-2000	2000US-0232080P
PR	08-SEP-2000	2000US-0232081P
PR	12-SEP-2000	2000US-0231968P
PR	14-SEP-2000	2000US-0233237P
PR	14-SEP-2000	2000US-0233238P
PR	14-SEP-2000	2000US-0233239P
PR	14-SEP-2000	2000US-0234423P
PR	21-SEP-2000	2000US-0234237P
PR	21-SEP-2000	2000US-0234238P
PR	21-SEP-2000	2000US-0234239P
PR	21-SEP-2000	2000US-0234240P
PR	21-SEP-2000	2000US-0234241P
PR	21-SEP-2000	2000US-0234242P
PR	21-SEP-2000	2000US-0234243P
PR	21-SEP-2000	2000US-0234244P
PR	21-SEP-2000	2000US-0234245P
PR	21-SEP-2000	2000US-0234246P
PR	21-SEP-2000	2000US-0234247P
PR	21-SEP-2000	2000US-0234248P
PR	21-SEP-2000	2000US-0234249P
PR	21-SEP-2000	2000US-0234250P
PR	21-SEP-2000	2000US-0234251P
PR	21-SEP-2000	2000US-0234252P
PR	21-SEP-2000	2000US-0234253P
PR	21-SEP-2000	2000US-0234254P
PR	21-SEP-2000	2000US-0234255P
PR	21-SEP-2000	2000US-0234256P
PR	21-SEP-2000	2000US-0234257P
PR	21-SEP-2000	2000US-0234258P
PR	21-SEP-2000	2000US-0234259P
PR	21-SEP-2000	2000US-0234260P
PR	21-SEP-2000	2000US-0234261P
PR	21-SEP-2000	2000US-0234262P
PR	21-SEP-2000	2000US-0234263P
PR	21-SEP-2000	2000US-0234264P
PR	21-SEP-2000	2000US-0234265P
PR	21-SEP-2000	2000US-0234266P
PR	21-SEP-2000	2000US-0234267P
PR	21-SEP-2000	2000US-0234268P
PR	21-SEP-2000	2000US-0234269P
PR	21-SEP-2000	2000US-0234270P
PR	21-SEP-2000	2000US-0234271P
PR	21-SEP-2000	2000US-0234272P
PR	21-SEP-2000	2000US-0234273P
PR	21-SEP-2000	2000US-0234274P
PR	21-SEP-2000	2000US-0234275P
PR	21-SEP-2000	2000US-0234276P
PR	21-SEP-2000	2000US-0234277P
PR	21-SEP-2000	2000US-0234278P
PR	21-SEP-2000	2000US-0234279P
PR	21-SEP-2000	2000US-0234280P
PR	21-SEP-2000	2000US-0234281P
PR	21-SEP-2000	2000US-0234282P
PR	21-SEP-2000	2000US-0234283P
PR	21-SEP-2000	2000US-0234284P
PR	21-SEP-2000	2000US-0234285P
PR	21-SEP-2000	2000US-0234286P
PR	21-SEP-2000	2000US-0234287P

PR	20-OCT-2000;	2000US-0241786P.
PR	20-OCT-2000;	2000US-0241787P.
PR	20-OCT-2000;	2000US-0241808P.
PR	20-OCT-2000;	2000US-0241809P.
PR	20-OCT-2000;	2000US-0241826P.
PR	01-NOV-2000;	2000US-0244617P.
PR	08-NOV-2000;	2000US-0246474P.
PR	08-NOV-2000;	2000US-0246475P.
PR	08-NOV-2000;	2000US-0246476P.
PR	08-NOV-2000;	2000US-0246477P.
PR	08-NOV-2000;	2000US-0246478P.
PR	08-NOV-2000;	2000US-0246522P.
PR	08-NOV-2000;	2000US-0246524P.
PR	08-NOV-2000;	2000US-0246525P.
PR	08-NOV-2000;	2000US-0246526P.
PR	08-NOV-2000;	2000US-0246527P.
PR	08-NOV-2000;	2000US-0246528P.
PR	08-NOV-2000;	2000US-0246532P.
PR	08-NOV-2000;	2000US-0246603P.
PR	08-NOV-2000;	2000US-0246610P.
PR	08-NOV-2000;	2000US-0246611P.
PR	08-NOV-2000;	2000US-0246613P.
PR	17-NOV-2000;	2000US-0249207P.
PR	17-NOV-2000;	2000US-0249208P.
PR	17-NOV-2000;	2000US-0249209P.
PR	17-NOV-2000;	2000US-0249210P.
PR	17-NOV-2000;	2000US-0249211P.
PR	17-NOV-2000;	2000US-0249212P.
PR	17-NOV-2000;	2000US-0249213P.
PR	17-NOV-2000;	2000US-0249214P.
PR	17-NOV-2000;	2000US-0249215P.
PR	17-NOV-2000;	2000US-0249216P.
PR	17-NOV-2000;	2000US-0249217P.
PR	17-NOV-2000;	2000US-0249218P.
PR	17-NOV-2000;	2000US-0249244P.
PR	17-NOV-2000;	2000US-0249245P.
PR	17-NOV-2000;	2000US-0249264P.
PR	17-NOV-2000;	2000US-0249265P.
PR	17-NOV-2000;	2000US-0249297P.
PR	17-NOV-2000;	2000US-0249299P.
PR	17-NOV-2000;	2000US-0249300P.
PR	01-DEC-2000;	2000US-0250160P.
PR	01-DEC-2000;	2000US-0250391P.
PR	05-DEC-2000;	2000US-0251030P.
PR	05-DEC-2000;	2000US-0251988P.
PR	05-DEC-2000;	2000US-0256719P.
PR	06-DEC-2000;	2000US-0251477P.
PR	08-DEC-2000;	2000US-0251856P.
PR	08-DEC-2000;	2000US-0251868P.
PR	08-DEC-2000;	2000US-0251869P.
PR	08-DEC-2000;	2000US-0251989P.
PR	08-DEC-2000;	2000US-0251990P.
PR	11-DEC-2000;	2000US-0254097P.
PR	05-JAN-2001;	2001US-0259678P.
XX	(HUMA-) HUMAN GENOME SCI INC.	
XX		
PI	Rosen CA, Barash SC, Ruben SM;	
XX	WPI; 2001-502630/55.	
DR		
XX		
PT	Polynucleotides encoding digestive system antigens, useful for	
PT	diagnosing, treating, preventing and/or prognosing disorders of the	
PT	digestive system, particularly cancer and cancer metastases.	
XX		
PS	Disclosure; SEQ ID NO 2638; 986bp; English.	
XX		
CC	The present invention provides the protein and coding sequences of a	
CC	number of human digestive system antigens. These can be used in the	
CC	diagnosis, treatment and prevention of digestive system disorders,	
CC	including cancer, Meckel's diverticulum, bacterial or parasitic	
CC	infections, appendicitis, Hirschsprung's disease, chronic colitis or	
CC	ulcerative colitis. The present sequence is a genomic DNA fragment	

CC encoding a digestive system antigen of the invention
XX
SQ Sequence 5468 BP; 1838 A; 981 C; 1052 G; 1597 T; 0 U; 0 Other;
Query Match 49.4%; Score 23.2; DB 4; Length 5468;
Best Local Similarity 83.3%; Pred. No. 3.4e+02;
Matches 25; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
OY 10 TGATTAGAAATAASTTTGATCACCATTTT 39
DB 1914 TGACTTGAAATAACTTGAACACAATTT 1885

Search completed: June 2, 2005, 08:57:27
Job time : 466 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2005, 08:20:08 / Search time 3128 Seconds
(without alignments)
571.938 Million cell updates/sec

Title: US-10-349-143-1

Perfect score: 47

Sequence: 1 ccacagcgtcgttagaatt.....gattcaccatttcaatttt 47

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

EST:*
1: gb_esc1:*
2: gb_esc2:*
3: gb_esc3:*
4: gb_esc4:*
5: gb_esc5:*
6: gb_esc6:*
7: gb_esc7:*
8: gb_esc8:*
9: gb_esc9:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27.8	59.1	463	8	AQ076343 CIT-HSP-2
2	27.6	58.7	945	9	CL473755 SAIL_205
3	26.4	56.2	1227	9	AG333342 Mus muscu
4	26.2	55.7	456	8	BX906921 Leishman
5	26.2	55.7	489	8	BZ196382 CH230-322
6	26	55.3	633	2	BB524921 BBS24921
7	26	55.3	2375	3	AK086348 Mus muscu
8	26	55.3	3058	3	AK043320 Mus muscu
9	25.8	54.9	349	1	AJ503513 AJ503513
10	25.8	54.9	507	1	AJ503688 AJ503688
11	25.6	54.5	765	9	CR071307 Reverse b
12	25.6	54.5	416	4	BI402346 MT-P-CPO-
13	25.6	54.5	422	7	CV294559 EST882936
14	25.6	54.5	443	7	CV300373 EST880889
15	25.6	54.5	453	7	CV294479 EST882856
16	25.6	54.5	462	7	CV293479 EST881856
17	25.6	54.5	493	7	CV293479 EST881856
18	25.6	54.5	495	7	CV300458 EST880974
19	25.6	54.5	507	7	CV300453 EST880969
20	25.6	54.5	540	7	CV295268 EST883645
21	25.6	54.5	555	7	CV294106 EST882483
22	25.6	54.5	614	7	CV297810 EST886187
23	25.6	54.5	620	7	CV293638 EST882015
24	25.6	54.5	630	7	CV294470 EST882847

25	25.6	54.5	638	7	CV293898	CV293898 EST882275
26	25.6	54.5	650	7	CV295618	CV295618 EST883995
27	25.6	54.5	658	7	CV297609	CV297609 EST885986
28	25.6	54.5	671	7	CV295246	CV295246 EST883623
29	25.6	54.5	673	7	CV298905	CV298905 EST887364
30	25.6	54.5	676	7	CV295713	CV295713 EST884090
31	25.6	54.5	755	5	BM417776	BM417776 UI-CF-EC1
32	25.4	54.0	758	5	BM417776	BM417776 UI-CF-EC1
33	25.4	54.0	1182	8	CC297513	CC297513 CH261-105
34	25.2	53.6	548	5	BP052517	BP052517 BP052517
35	25.2	53.6	609	1	AL655320	AL655320 AL655320
36	25.2	53.6	836	8	BZ169970	BZ169970 CH230-414
37	25	53.2	607	8	BZ169970	BZ169970 CH230-414
38	25	53.2	1101	9	CNSG00FRJ	AL071159 Drosophila
39	24.8	52.8	272	2	BB338231	BB338231 BB338231
40	24.8	52.8	488	8	B41058	B41058 HS-1052-B2-
41	24.8	52.8	521	5	BP181191	BP181191 BP181191
42	24.8	52.8	601	7	CNS33313	CNS33313 UI-M-HO-
43	24.8	52.8	639	2	BB252497	BB252497 BB252497
44	24.8	52.8	702	8	BH583357	BH583357 odg79H04.
45	24.8	52.8	707	8	BH551295	BH551295 B0H6256TR
46	24.8	52.8	746	8	CC118714	CC118714 NDL-58C8.
47	24.8	52.8	766	8	BH577556	BH577556 B0GH158TF
48	24.8	52.8	773	8	BH498017	BH498017 B0GJX32TF
49	24.8	52.8	916	8	AZ533670	AZ533670 ENTB150TR
50	24.8	52.8	1072	9	CL021677	CL021677 CH216-8L2
51	24.8	52.8	3367	3	AK043061	AK043061 Mus muscu
52	24.6	52.3	470	6	BY600070	BY600070 BY600070
53	24.6	52.3	498	8	AQ370013	AQ370013 HS-5045-A
54	24.6	52.3	504	5	BQ487419	BQ487419 WHE1713-1
55	24.6	52.3	816	8	AZ157787	AZ157787 SP-0058-A
56	24.6	52.3	831	9	BX180649	BX180649 Dario rct
57	24.4	51.9	132	7	CNS93796	CNS93796 EC2BBA8CH
58	24.4	51.9	225	9	CNS98894	CNS98894 CH240-145
59	24.4	51.9	475	1	AA768788	AA768788 OC17909-8
60	24.4	51.9	493	5	BM557140	BM557140 BM557140
61	24.4	51.9	635	6	CB507009	CB507009 B8A10B505
62	24.4	51.9	662	5	BQ273840	BQ273840 rc71d01.Y
63	24.4	51.9	702	8	BZ939750	BZ939750 CH240-106
64	24.4	51.9	725	7	CNS89138	CNS89138 EC2BBA31A
65	24.4	51.9	747	6	CA063687	CA063687 B8A11G5B3
66	24.2	51.5	198	8	AZ263847	AZ263847 RPCI-23-1
67	24.2	51.5	587	6	BY691928	BY691928 BY691928
68	24.2	51.5	443	9	CNS068163	CNS068163 Reverse b
69	24.2	51.5	450	5	BX954363	BX954363 DKFZP781P
70	24.2	51.5	474	8	AZ262790	AZ262790 RPCI-23-1
71	24.2	51.5	556	9	CL592069	CL592069 OB-Ba005
72	24.2	51.5	587	8	BH199785	BH199785 Sml-39A11
73	24.2	51.5	589	9	CE205000	CE205000 t1gr-g88-
74	24.2	51.5	657	5	BH433553	BH433553 603219945
75	24.2	51.5	708	8	AZ870628	AZ870628 ZK0183K12
76	24.2	51.5	879	9	CL133935	CL133935 ISB1-104D
77	24.2	51.5	946	5	BH849838	BH849838 AGENCOURT
78	24.2	51.5	972	8	CC085279	CC085279 GCU-K33r.
79	24.2	51.5	1122	9	CL649974	CL649974 CH213-264
80	24.2	51.5	1170	6	CD507583	CD507583 CNA86-A06
81	24.2	51.5	5876	3	CR749668	CR749668 Homo sapi
82	24	51.1	483	7	CV293480	CV293480 EST881857
83	24	51.1	583	5	BP334038	BP334038 BP334038
84	24	51.1	621	8	BH037779	BH037779 RPCI-24-3
85	24	51.1	647	8	AO384267	AO384267 RPCI11-13
86	24	51.1	734	6	CB634378	CB634378 OS11EB13N
87	24	51.1	875	8	CC000427	CC000427 PDBB14TD
88	24	51.1	1220	8	CC205996	CC205996 CH261-56U
89	23.8	50.6	360	7	D18433	D18433 MUCSG01491
90	23.8	50.6	377	7	CK150865	CK150865 SWMB13CAW
91	23.8	50.6	420	7	CK150865	CK150865 CMPEST02
92	23.8	50.6	431	6	BY615187	BY615187 BY615187
93	23.8	50.6	471	9	CL349441	CL349441 RPCI14-P00
94	23.8	50.6	477	7	CK703544	CK703544 ZP101-P30
95	23.8	50.6	500	5	BM571363	BM571363 BM571363
96	23.8	50.6	513	8	BH760753	BH760753 BMBAC316F

98 23.8 50.6 533 8 A0524290 HS_5237_B
99 23.8 50.6 548 4 B0692271 B0692271
100 23.8 50.6 554 4 BM31372 BM31372 PIC1_65_C

ALIGNMENTS

RESULT 1
A0076343 463 bp DNA linear GSS 20-AUG-1998
LOCUS CIT-HSP-2359G24.TF CIT-HSP Homo sapiens genomic clone 2359G24,
DEFINITION genomic survey sequence.
ACCESSION A0076343
VERSION A0076343.1 GI:3437527
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 463)
AUTHORS Adams,M.D., Rounley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
Venter,J.C.
TITLE Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
JOURNAL Unpublished (1998)
COMMENT Other GSSs: CIT-HSP-2359G24.TR
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/cdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.

FEATURES

source location/Qualifiers
1..463
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="2359G24"
/sex="Male"
/cell_type="Sperm"
/clone_lib="CIT-HSP"
/note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"

ORIGIN

Query Match 59.1%; Score 27.8; DB 8; Length 463;
Best Local Similarity 78.0%; Pred.No.1.2e+02;
Matches 32; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Oy 6 GCGCTGATTAGAAATTAATGATCACCATTTCGAATTT 46
||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 GCGCTTGTAGTAAACATGATTAACATTTTAAATTT 461

RESULT 2
LOCUS CL473755/c 945 bp DNA linear GSS 01-APR-2004
DEFINITION SAIL_205.E07.v1 SAIL Collection Arabidopsis thaliana genomic clone
CL473755-
SAIL_205.E07.v1, genomic survey sequence.
ACCESSION CL473755
VERSION CL473755.1 GI:45939466
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (chale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE
AUTHORS Sessions,A., Burke,E., Presting,G., Aux,G., McElver,J., Patton,D.,
1 (bases 1 to 945)
Sessions,A., Burke,E., Presting,G., Aux,G., Clarke,J.D., Cotton,D.,
Bullis,D., Snell,J., Miguel,T., Hutchison,D., Kimerly,B.,
Mitxel,T., Katsagiri,F., Glazedbrook,J., Law,M. and Goff,S.A.
A high-throughput Arabidopsis reverse genetics system
Plant Cell 14 (12), 2985-2994 (2002)
22356987
MEDLINE
PUBMED 12468722

COMMENT

Contract: Sessions A
Applied Trait Genetics
Syngenta Biotechnology Inc.
3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA
Email: allen.sessions@syngenta.com
ABRC Stock Number CS809613; T-DNA left border flanking sequences of
Syngenta Arabidopsis Insertion Library (SAIL) lines are available
through the Arabidopsis Biological Resource Center (ABRC).
Sequences represent a pool of amplified genomic regions and not
single contiguous sequences.
Class: TDNA tagged.
location/Qualifiers
1..945
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="SAIL_205.E07.v1"
/clone_lib="SAIL Collection"
/note="T-DNA left border sequences were isolated using a
modified Tail-PCR strategy"

FEATURES

source

ORIGIN

Query Match 58.7%; Score 27.6; DB 9; Length 945;
Best Local Similarity 75.0%; Pred.No.1.4e+02;
Matches 33; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

Oy 4 CAGCGCTGATTAGAAATTAATGATCACCATTTCGAATTT 47
||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 375 CAGCGTTGTAGCAATAGATTGATCTTGATTTAGAACTTT 332

RESULT 3
LOCUS AG333342 1227 bp DNA linear GSS 02-JUN-2004
DEFINITION Mus musculus molossinus DNA, clone:MSMg01-123N09.TU, genomic survey
sequence.
ACCESSION AG333342
VERSION AG333342.1 GI:47906652
KEYWORDS GSS.
SOURCE Mus musculus molossinus
ORGANISM Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
BAC End Sequences of Library MSMg01
2 (bases 1 to 1227)
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
Direct Submission
Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsukuba-shi, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@psc.riken.jp, URL:http://ngp.gsc.riken.go.jp/
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center.
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199

COMMENT

e-mail: abe@rtc.riken.jp

```
Sequencing : TJ
LIBRARY
Vector      : pBACe3.6
R.Site 1    : EcoRI
R.Site 2    : EcoRI.
```

FEATURES	Location/Qualifiers
SOURCE	1. .1227

```

/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-123N09.7J"
/sex="male"
/tissue_type="mixture of kidney and spleen"
clone_lib="MSMg01 Mouse Male BAC library"

```

Query Match	56.2%	Score 26.4	DB 9	Length 1227
Best Local Similarity	71.7%	Pred. No. 3.5e+02		
Matches 33, Conservative	1	Mismatches 12	Indels 0	Gaps 0

```

QY      2  CACAGCGTTGATTAGAAATAASTTTGATCACCATTTCAAATTTT  47
        ||||| ||||| : ||| ||||| ||||| |||||
DB      98  CACAAGCTTACTTAGAAGTACCCTCTGTGTACACATCTTGAATTTT  143

```

LOCUS	456 bp	DNA	linear	GSS 07-JAN-2004
DEFINITION	BX9065921			
	Leishmania braziliensis	GSS, clone	LBAF74F03,	genomic survey
	sequence.			

VERSION BK906921.1 GI:40735369
KEYWORDS GSS; genomic survey sequence.
SOURCE Leishmania braziliensis
ORGANISM Leishmania braziliensis

REFERENCE
1
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL
2
2 (bases 1 to 456)
Cruz, A.K.
Direct Submission
Submitted (07-JUN-2004) Cruz A.K., University of S

source	1. .456
--------	---------

ORIGIN

Query Match	55.7%	Score 26.2	DB 9	Length 456
Best Local Similarity	79.5%	Pred. No. 3.9e+02		
Matches 31, Conservative	0	Mismatches 8	Indels 0	Gaps 0

```

QY      1  CCACAGCTGATTAGAAATAASTTTGATCACCATTTT  35
          |||||  |||  |||||  |||||  |||||
DB      127 CCACAGCATGAAAAAAGTAAATTTGACCCCATTTT  89

```

RESULT 5	BZ196382	489 bp	DNA	linear	GSS 11-OCT-2002
LOCUS	BZ196382				
DEFINITION	CH230-322A14.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone				

ACCESSION	BZ196382
VERSION	BZ196382.1
KEYWORDS	GI:23854434
SOURCE	GSS.
	<i>Rattus norvegicus</i> (Norway rat)

SOURCE ORGANISM	REFERENCE	TITLE	JOURNAL	COMMENT
Rattus norvegicus (Norway rat)	1 (bases 1 to 489)			
Rattus norvegicus	Zhao, S., Shetty, J., Shateman, S., Teegaye, G., Geer, K., Shwartzsbeym, A., Gelregeorgis, E., Overton, L., Riggs, F., de Jong, P. and Fraser, C. M.			
Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Eueleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus		Rat BAC and Sequences from Library CHORI-230 MboI segment Unpublished (1989)		
	Other_GSSs: CH230-322A14..TV			

FEATURES	Location/Qualifiers
source	1. .489

```

organism="Rattus norvegicus"
mol_type="genomic DNA"
strain="BN/SsNHsd/MCM"
db_xref="taxon:10116"
clone="CH230-322A14"
sex="Female"
cell_type="Brain"
clone_lib="CHORI-230 Segment 2"
note="Vector: pIRABK1.3; Site_1: MboI, Site_2: MboI
CHORI-230 Rat (BN/SsNHsd/MCM) BAC library produced by
Pleter de Jong"

```

ORIGIN

Query Match	55.7%	Score 26.2;	DB 8;	Length 489;
Best Local Similarity	75.6%	Pred. No. 3.9e+02;		
Matches 31; Conservative	1;	Mismatches 9;	Indels 0;	Gaps 0;

```

QY      7  GCTTGATTAGAAATAAASTTTGATCACCATTTCCAAATTTT  47
          ||| ||| ||| : ||| ||| ||| ||| ||| ||| |||
Db    385 GCATGGTTGGATCCACTTGTGATCACCATTTCATATCTT  42

```

LOCUS	633 bp	mRNA	linear	EST 25-OCT-2001
BB524821	RIKEN full-length enriched, 15 days embryo			head Mus
BB524821	musculus cDNA D930023E13 3', mRNA sequence.			

VERSION	BB524921.2	GI:16444057
KEYWORDS	EST.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	

REFERENCE
AUTHORS

Bukayeva, Meltzoi, Chordata; Craniata; Vertebrata; Euteleostomi;
Eukarya; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
1 (bases 1 to 633)

Akaiawa, Y., Carnivora, P., Fukuda, S., Furuno, M., Hasegaki, T.,
Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kanai, J.,
Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K.,
Onno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K.

TITLE
JOURNAL

Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

FEATURES
source

location/Qualifiers
1..2375
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:D930023E13"
/db_xref="taxon:10090"
/clone="D930023E13"
/tissue_type="head"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="15 days embryo"
688..987
/note="unnamed protein product; 5-HYDROXYTRYPTAMINE 2C RECEPTOR [SWISSPROT|P34968, evidence: FASTA, 99.6%ID, 100%length, match=1377)
putative"
/codon_start=1
/protein_id="BAC39653.1"
/db_xref="GI:26352033"
/translation="MVNIGTRVRSLSLVLHLGLWQPIGISPSPAAYTPFNSSDGSGRLFPPGVQMPLSLIVIIIMITGGIILVIWASMEKKLNATNYFLMSLGHC"

ORIGIN

Query Match 55.3%; Score 26; DB 3; Length 2375;
Beet Local Similarity 80.6%; Pred. No. 4.8e+02;

Matches 29; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Df Db 2231 ATTAGCAATTAACTTGTAACCTTTGAATTTCAAATTTT 2196

Cy |||||
|||||

AKO43320/c AKO43320 3058 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus 7 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:A730084N2 Product:5-HYDROXYTRYPTAMINE 2C RECEPTOR, Full insert sequence.
ACCESSION AKO43320
VERSION AKO43320.1 GI:26089632
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Euarchyotae; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Carninci,P., Shibata,Y., Hayase,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okasaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PubMed 10349636

REFERENCE
AUTHORS Carninci,P., Shibata,Y., Hayase,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okasaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL MEDLINE PUBMED AUTHORS	Genome Res. 10 (10), 1617-1630 (2000) 20499374 11042159
JOURNAL MEDLINE PUBMED AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Komu,H., Akiyama,J., Nishi,K., Kitanai,T., Tashiro,H., Itoh,M., Suni,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A., Yamamoto,R., Matsunoto,H., Sakaguchi,S., Ikegami,T., Kasaiyagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multipipillary sequencer Genome Res. 10 (11), 1757-1771 (2000) 20530913 11076861
JOURNAL MEDLINE PUBMED AUTHORS	4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001) 5
JOURNAL MEDLINE PUBMED AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002) 6 (bases 1 to 3058) Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,M., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiroka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Komu,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohashi,T., Ohtsuki,Y., Saito,R., Satoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaki-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y. Direct Submission
JOURNAL MEDLINE PUBMED AUTHORS	Submitted (15-JUL-2001) Yoshihide Hayashizaki. The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp). URL: http://genome.gsc.riken.jp/ , Tel: 81-45-503-9222, Fax: 81-45-503-9216 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ , URL: http://fantom.gsc.riken.jp/ , Location/Qualifiers
FEATURES	1..3058 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="PANTOM.DB:A730084N12" /db_xref="taxon:10090" /clone="A730084N12" /clone_type="cerebellum" /clone_id="RIKEN full-length enriched mouse cDNA library" /dev_stage="7 days neonate" 1..3058 /note="5-HYDROXYTRYPTAMINE 2C RECEPTOR [SWISSPROT P34968, evidence:FASTA, 99.6%ID, 100%length, match=1377]"
ORIGIN	misc_feature
Query Match	55.3%; Score 26; DB 3; Length 3058;

Best Local Similarity 80.6%; Pred. No. 4.9e+02;
Matches 29; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 12 ATTGAAATTAATTTGATCACCATTTTCAAAATTTT 47
Db 2229 ATTGCAATTAACCTTGAACTTGATTTTCAAAATTTT 2194

RESULT 9

AJ503513

LOCUS AJ503513 MTAMP Medicago truncatula cDNA clone mtgmadc120035a04, 349 bp mRNA linear EST 10-FEB-2003

DEFINITION mRNA sequence.

ACCESSION AJ503513 GI:22084441

VERSION AJ503513.1 GI:22084441

KEYWORDS EST.

SOURCE Medicago truncatula (barrel medic)

ORGANISM Medicago truncatula

REFERENCE 1 (bases 1 to 349)

AUTHORS Manthey,K., Bartelsmeier,V., Baier,M.C., Meyer,F., Bartels,D.,

TITLE Bekel,T., Linke,B., Grunwald,U., Franken,P., Kuester,H.,

JOURNAL Perlick,A.M. and Puehler,A.

COMMENT Detection of transcript sequences from mycorrhizal roots of the

model mycorrhiza Medicago truncatula genotype A17 - Glomus

intraradices using the approach of an EST genome project

Unpublished (2002)

CONTACT: Kuester H

Lehrstuhl fuer Genetik

Universitaet Bielefeld

Postfach 100131, D-33501 Bielefeld, Germany.

FEATURES

source

1..349

/organism="Medicago truncatula"

/mol_type="mRNA"

/db_xref="taxon:3880"

/clone="mtgmadc120035a04"

/tissue_type="mycorrhizal roots"

/dev_stage="six week old mycorrhizal roots harvested five

weeks after inoculation with Glomus intraradices"

/clone_1ib="MTAMP"

/note="Vector: pGEM-T; Site 1: PstI; Site 2: SphI;

genotype A17; cDNA was prepared from polyA+ enriched RNA

from mycorrhizal roots harvested five weeks after

inoculation. The cDNA was directionally ligated by

Medigenomix into the pGEM-T vector from Promega using

GCATCGCGCGAGCGCGCGACATG and CTGCAAGCCATTATGCGCGG

adapters. Plasmids containing cDNA inserts were propagated

in E. coli DH10B cells."

ORIGIN

Query Match 54.9%; Score 25.8; DB 1; Length 349;

Best Local Similarity 76.9%; Pred. No. 5.1e+02;

Matches 30; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 5 AGCCTTGATTAGAAATTAATTTGATCACCATTTTCAAA 43

Db 71 AGGCTTCATTAATAATTCACCTTTATTCACCATTTTCAAA 109

RESULT 10

AJ503688

LOCUS AJ503688 MTAMP Medicago truncatula cDNA clone mtgmadc120037a12, 507 bp mRNA linear EST 10-FEB-2003

DEFINITION mRNA sequence.

ACCESSION AJ503688

VERSION AJ503688.1 GI:22084616

KEYWORDS EST.

SOURCE Medicago truncatula (barrel medic)

ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

REFERENCE

1 (bases 1 to 507)

AUTHORS Manthey,K., Bartelsmeier,V., Baier,M.C., Meyer,F., Bartels,D.,

TITLE Bekel,T., Linke,B., Grunwald,U., Franken,P., Kuester,H.,

JOURNAL Perlick,A.M. and Puehler,A.

COMMENT Detection of transcript sequences from mycorrhizal roots of the

model mycorrhiza Medicago truncatula genotype A17 - Glomus

intraradices using the approach of an EST genome project

Unpublished (2002)

CONTACT: Kuester H

Lehrstuhl fuer Genetik

Universitaet Bielefeld

Postfach 100131, D-33501 Bielefeld, Germany.

FEATURES

source

1..507

/organism="Medicago truncatula"

/mol_type="mRNA"

/db_xref="taxon:3880"

/clone="mtgmadc120037a12"

/tissue_type="mycorrhizal roots"

/dev_stage="six week old mycorrhizal roots harvested five

weeks after inoculation with Glomus intraradices"

/clone_1ib="MTAMP"

/note="Vector: pGEM-T; Site 1: PstI; Site 2: SphI;

genotype A17; cDNA was prepared from polyA+ enriched RNA

from mycorrhizal roots harvested five weeks after

inoculation. The cDNA was directionally ligated by

Medigenomix into the pGEM-T vector from Promega using

GCATCGCGCGAGCGCGCGACATG and CTGCAAGCCATTATGCGCGG

adapters. Plasmids containing cDNA inserts were propagated

in E. coli DH10B cells."

ORIGIN

Query Match 54.9%; Score 25.8; DB 1; Length 507;

Best Local Similarity 76.9%; Pred. No. 5.2e+02;

Matches 30; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 5 AGCCTTGATTAGAAATTAATTTGATCACCATTTTCAAA 43

Db 71 AGGCTTCATTAATAATTCACCTTTATTCACCATTTTCAAA 109

RESULT 11

CR071307

LOCUS CR071307/c 765 bp DNA linear GSS 05-JUL-2004

DEFINITION Reverse strand read from insert in 5'HPR insertion targeting and

chromosome engineering clone MHPN201114, genomic survey sequence.

CR071307.1 GI:49804897

VERSION GSS; genome survey sequence; MICER.

KEYWORDS GSS; musculus (house mouse)

SOURCE Mus musculus

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 765)

AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,

Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,

Rogers,J. and Bradley,A.

TITLE Direct Submission

JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. http://www.sanger.ac.uk/MICER

FEATURES

source

1..765

/organism="Mus musculus"

/mol_type="Genomic DNA"

/db_xref="taxon:10090"

/clone="MHPN201114"

/clone_1ib="MHPN"

ORIGIN

Query Match 54.9%; Score 25.8; DB 9; Length 765;
 Best Local Similarity 76.9%; Pred. No. 5.3e+02;
 Matches 30; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 5 AGGCTTGATTAGAAATTAAGTTGATCACCATTTCAAA 43
 DB 716 AAGCTTGAAATAGAGATAAGTTGGAGACTATTTTACA 678

RESULT 12 BI402346 416 bp mRNA linear EST 14-AUG-2001
 LOCUS BI402346/c MI-P-CPD-nw-a-03-0-UI.61 MI-P-CPD Sus scrofa cDNA clone
 DEFINITION MI-P-CPD-nw-a-03-0-UI 3', mRNA sequence.
 ACCESSION BI402346.1 GI:15181407
 VERSION BI402346.1
 KEYWORDS EST.
 SOURCE Sus scrofa (pig)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1 (bases 1 to 416)
 AUTHORS Bonaldo,M.P., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548

COMMENT Contact: Tugile CK
 Molecular Genetics Laboratory, Department of Animal Science
 Iowa State University
 201 Kildee Hall, Ames, IA 50011-3150, USA
 Tel: 5152944252
 Fax: 5152942401
 Email: Cktugile@iastate.edu

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to verify it as a clone from the non-normalized uterine library cDNA library Preparation: M.B. Soares Lab, University of Iowa EST sequencing: M.B. Soares Lab, University of Iowa Clone distribution: clones will be available through Research Genetics (www.reagen.com) The following repetitive elements were found in this cDNA sequence: 1-24,
 >AT rich#low complexity
 Seq primer: M13 Forward
 POLYA=Yes.

FEATURES

source Location/Qualifiers
 1..416
 /organism="Sus scrofa"
 /mol_type="mRNA"
 /db_xref="taxon:9823"
 /clone="MI-P-CPD-nw-a-03-0-UI"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="MI-P-CPD"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: EcoRI; The MI-P-CPD library is derived from uteris. For a detailed description of the library from which this clone was derived, please visit our web site at <http://piggen.genome.iastate.edu/>. The procedure used to create this library has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)
 TAG_TISSUE=uterus
 TAG_LIB=MI-P-CPD
 TAG_SEQ=AGTCCATCG"

ORIGIN

Query Match 54.5%; Score 25.6; DB 4; Length 416;
 Best Local Similarity 82.4%; Pred. No. 6e+02;
 Matches 28; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 10 TGATTAGAAATTAAGTTGATCACCATTTCAAA 43
 DB 48 TGACTGAAATTAAGCTTGACACCAATTTTAAA 15

RESULT 13 CV294559 422 bp mRNA linear EST 23-SEP-2004
 LOCUS CV294559/c EST862936 petunia floral post-ethylene cDNA library Petunia x hybrida cDNA clone Petunia-C2H4-20-G12 5' end, mRNA sequence.
 DEFINITION hybrida cDNA clone Petunia-C2H4-20-G12 5' end, mRNA sequence.
 ACCESSION CV294559
 VERSION CV294559.1 GI:52583969
 KEYWORDS EST.
 SOURCE Petunia x hybrida
 ORGANISM Petunia x hybrida
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asteridae; lamiales; Solanales; Solanaceae; Petunia.

REFERENCE 1 (bases 1 to 422)
 AUTHORS Shibusaki,K., Underwood,B., Loucas,H., Farnerie,W., Jones,M. and Clark,D.
 TITLE Petunia x hybrida EST collection
 JOURNAL Unpublished (2004)
 COMMENT Contact: David Clark
 UF floriculture Biotechnology Lab
 University of Florida
 Environmental Horticulture Department, 1545 Fifield Hall, Box 110670, Gainesville, FL 32611-0670, USA
 Tel: 352-392-1831 x370
 Fax: 352-392-3870
 Email: dclark@mail.ifas.ufl.edu
 Contact Dr. Clark (dclark@mail.ifas.ufl.edu) for clone information
 Seq primer: T3 primer.

FEATURES

source Location/Qualifiers
 1..422
 /organism="Petunia x hybrida"
 /mol_type="mRNA"
 /cultivar="Mitchell Diploid (aka. Mitchell, aka W15 in Europe)"
 /db_xref="taxon:4102"
 /clone="Petunia-C2H4-20-G12"
 /tissue_type="all floral organs"
 /lab_host="lambda ZAPI unidirectional"
 /clone_lib="petunia floral post-ethylene cDNA library"
 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; supplier: Petunia x hybrida cv. Mitchell Diploid plants were grown from seeds to a fully flowering stage under standard greenhouse conditions. Flowers at anthesis stage were excised from plants grown in standard greenhouses and treated with 2ppm exogenous ethylene. Entire flowers were collected at 0, 5, 10, 15, 24, and 36 hours after treatment. Total RNA was extracted from each sample, and 100 micrograms of each sample was combined for subsequent poly A+ mRNA selection and cDNA synthesis."

ORIGIN

Query Match 54.5%; Score 25.6; DB 7; Length 422;
 Best Local Similarity 73.8%; Pred. No. 6e+02;
 Matches 31; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 6 GGCTTGATTAGAAATTAAGTTGATCACCATTTCAAATTTT 47
 DB 327 GGCTGCATCAACAATAATGTTGATCACCCTTCACAATGCT 286

RESULT 14 CV300373 443 bp mRNA linear EST 23-SEP-2004
 LOCUS CV300373/c EST860889 petunia floral development cDNA library Petunia x hybrida
 DEFINITION cDNA clone Petunia-tesq1-D04 5' end, mRNA sequence.
 ACCESSION CV300373
 VERSION CV300373.1 GI:52595614
 KEYWORDS EST.

Query Match 54.5%; Score 25.6; DB 7; Length 422;
 Best Local Similarity 73.8%; Pred. No. 6e+02;
 Matches 31; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

Petunia x hybrida
Petunia x hybrida
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Petunia.
1 (bases 1 to 443)
Shibuya, K., Underwood, B., Loucas, H., Farmerie, W., Jones, M. and
Clark, D.
Petunia x hybrida EST collection
Unpublished (2004)
Contact: David Clark
UP floriculture Biotechnology Lab
University of Florida
Environmental Horticulture Department, 1545 Field Hall, Box
110670, Gainesville, FL 32611-0670, USA
Tel: 352-392-1831 x370
Fax: 352-392-3870
Email: dclark@mail.ifas.ufl.edu
Contact Dr. Clark (dclark@mail.ifas.ufl.edu) for clone information
Seq primer: T3 primer.
Location/Qualifiers
1. 443
/organism="Petunia x hybrida"
/mol_type="mRNA"
/cultivar="Mitchell Diploid (aka. Mitchell, aka W15 in
Europe)"
/db_xref="taxon:4102"
/clone="Petunia-iresql-D04"
/tissue_type="all floral organs"
/lab_host="Lambda ZAPII unidirectional"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; supplier: Petunia x hybrida cv. Mitchell Diploid
plants were grown from seeds to a fully flowering stage
under standard greenhouse conditions. Ten entire flower
of six developmental stages were collected on the same day
from plants grown in standard greenhouses. The flower
stages were as follows in chronological order from
youngest to oldest: stage 1 - no color in corolla; corolla
0.5 inches long stage 2 - first sign of color in corolla;
corolla .75-1 inches long stage 3 - fully elongated
corolla (not open); corolla 1.5 inches long stage 4 -
fully open corolla; anthers not yet dehiscent stage 5 -
pollen; wet stigma stage 6 - pre-senescent; yellowing of
corolla tube; dry brown pollen (if present); stigma dry.
Total RNA was extracted from each sample, and 100
micrograms of each sample was combined for subsequent poly
A+ mRNA selection and cDNA synthesis."

ORIGIN

Query Match 54.5%; Score 25.6; DB 7; Length 443;
Best Local Similarity 73.8%; Pred. No. 6e+02; Mismatches 10; Indels 0; Gaps 0;
Matches 31; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

Qy 6 GGCTTGATTAGAAATTAATTTGATCACCATTTCGAATTTT 47
||||| :|||||
Db 257 GGCTGATCACAATAATTGATTCACCTTCACAAATGTT 216
||||| :|||||

RESULT 15
CV294479/c 453 bp mRNA linear EST 23-SEP-2004
LOCUS
DEFINITION EST882856 petunia floral post-ethylene cDNA library Petunia x
hybrida cDNA clone Petunia-C2H4-20-A01 5' end, mRNA sequence.
ACCESSION
CV294479
VERSION
CV294479.1 GI:52583811
KEYWORDS
EST.
SOURCE
ORGANISM
Petunia x hybrida
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Petunia.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

1 (bases 1 to 453)
Shibuya, K., Underwood, B., Loucas, H., Farmerie, W., Jones, M. and
Clark, D.
Petunia x hybrida EST collection
Unpublished (2004)
Contact: David Clark
UP floriculture Biotechnology Lab
University of Florida
Environmental Horticulture Department, 1545 Field Hall, Box
110670, Gainesville, FL 32611-0670, USA
Tel: 352-392-1831 x370
Fax: 352-392-3870
Email: dclark@mail.ifas.ufl.edu
Contact Dr. Clark (dclark@mail.ifas.ufl.edu) for clone information
Seq primer: T3 primer.
Location/Qualifiers
1. 453
/organism="Petunia x hybrida"
/mol_type="mRNA"
/cultivar="Mitchell Diploid (aka. Mitchell, aka W15 in
Europe)"
/db_xref="taxon:4102"
/clone="Petunia-C2H4-20-A01"
/tissue_type="all floral organs"
/lab_host="Lambda ZAPII unidirectional"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; supplier: Petunia x hybrida cv. Mitchell Diploid
plants were grown from seeds to a fully flowering stage
under standard greenhouse conditions. Flowers at anthesis
stage were excised from plants grown in standard
greenhouses and treated with 2ppm exogenous ethylene.
Entire flowers were collected at 0, 5, 10, 15, 24, and 36
hours after treatment. Total RNA was extracted from each
sample, and 100 micrograms of each sample was combined for
subsequent poly A+ mRNA selection and cDNA synthesis."

ORIGIN

Query Match 54.5%; Score 25.6; DB 7; Length 453;
Best Local Similarity 73.8%; Pred. No. 6e+02; Mismatches 10; Indels 0; Gaps 0;
Matches 31; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

Qy 6 GGCTTGATTAGAAATTAATTTGATCACCATTTCGAATTTT 47
||||| :|||||
Db 327 GGCTGATCACAATAATTGATTCACCTTCACAAATGTT 286
||||| :|||||

RESULT 16
CV295114 462 bp mRNA linear EST 23-SEP-2004
LOCUS
DEFINITION EST883491 petunia floral post-ethylene cDNA library Petunia x
hybrida cDNA clone Petunia-C2H4-25-A04 5' end, mRNA sequence.
ACCESSION
CV295114
VERSION
CV295114.1 GI:52585086
KEYWORDS
EST.
SOURCE
ORGANISM
Petunia x hybrida
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Petunia.
1 (bases 1 to 462)
Shibuya, K., Underwood, B., Loucas, H., Farmerie, W., Jones, M. and
Clark, D.
Petunia x hybrida EST collection
Unpublished (2004)
Contact: David Clark
UP floriculture Biotechnology Lab
University of Florida
Environmental Horticulture Department, 1545 Field Hall, Box
110670, Gainesville, FL 32611-0670, USA
Tel: 352-392-1831 x370
Fax: 352-392-3870
Email: dclark@mail.ifas.ufl.edu

Contact Dr. Clark (dclark@mail.iffas.ufl.edu) for clone information
Seq primer: T3 primer.

FEATURES

Location/Qualifiers
1..462

/organism="Petunia x hybrida"
/mol_type="mRNA"
/cultivar="Mitchell Diploid (aka. Mitchell, aka W15 in Europe)"
/db_xref="taxon:4102"
/clone="Petunia-C2H4-29-A04"
/tissue_type="all floral organs"
/lab_host="lambda ZAPII unidirectional"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; supplier: Petunia x hybrida cv. Mitchell Diploid plants were grown from seeds to a fully flowering stage under standard greenhouse conditions. Flowers at anthesis stage were excised from plants grown in standard greenhouses and treated with 2ppm exogenous ethylene. Entire flowers were collected at 0, 5, 10, 15, 24, and 36 hours after treatment. Total RNA was extracted from each sample, and 100 micrograms of each sample was combined for subsequent poly A+ mRNA selection and cDNA synthesis."

ORIGIN

Query Match 54.5%; Score 25.6; DB 7; Length 462;
Best Local Similarity 73.8%; Pred. No. 6e+02;
Matches 31; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

Qy 6 GGCTGATTAGAAATTAAGTTGATCACCATTTCATTTT 47
|||||
Db 324 GGCTGCATCACAATAATTAGTTGATCACCCTTCACAAATGTT 283
|||||

RESULT 17 493 bp mRNA linear EST 23-SEP-2004
CV293479 petunia floral post-ethylene cDNA library Petunia x
LOCUS EST81856 petunia floral post-ethylene cDNA library Petunia x
DEFINITION hybrida cDNA clone Petunia-C2H4-6R-F05 5' end, mRNA sequence.
ACCESSION CV293479
VERSION CV293479.1 GI:52581808
KEYWORDS EST.
SOURCE Petunia x hybrida
ORGANISM Petunia x hybrida
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Petunia.
1 (bases 1 to 493)
Tel: 352-392-1831 x370
Fax: 352-392-3870
Email: dclark@mail.iffas.ufl.edu
Contact Dr. Clark (dclark@mail.iffas.ufl.edu) for clone information
Seq primer: T3 primer.

REFERENCE AUTHORS
Clark, D., Underwood, B., Loucas, H., Farmerie, W., Jones, M. and
Petunia x hybrida EST collection

TITLE Unpublished (2004)
JOURNAL Contact: David Clark
COMMENT UF Horticulture Biotechnology Lab
University of Florida
Environmental Horticulture Department, 1545 Fifield Hall, Box
110670, Gainesville, FL 32611-0670, USA
Tel: 352-392-3870
Fax: 352-392-3870
Email: dclark@mail.iffas.ufl.edu
Contact Dr. Clark (dclark@mail.iffas.ufl.edu) for clone information
Seq primer: T3 primer.

FEATURES

source

1..493
Location/Qualifiers
/organism="Petunia x hybrida"
/mol_type="mRNA"
/cultivar="Mitchell Diploid (aka. Mitchell, aka W15 in Europe)"
/db_xref="taxon:4102"
/clone="Petunia-C2H4-6R-F05"
/tissue_type="all floral organs"
/lab_host="lambda ZAPII unidirectional"
/clone_lib="petunia floral post-ethylene cDNA library"

ORIGIN

Query Match 54.5%; Score 25.6; DB 7; Length 493;
Best Local Similarity 73.8%; Pred. No. 6e+02;
Matches 31; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

Qy 6 GGCTGATTAGAAATTAAGTTGATCACCATTTCATTTT 47
|||||
Db 201 GGCTGCATCACAATAATTAGTTGATCACCCTTCACAAATGTT 160
|||||

RESULT 18 495 bp mRNA linear EST 23-SEP-2004
CV300458 petunia floral development cDNA library Petunia x hybrida
LOCUS EST860974 petunia floral development cDNA library Petunia x hybrida
DEFINITION cDNA clone Petunia-resq1Teet-D12 5' end, mRNA sequence.
ACCESSION CV300458
VERSION CV300458.1 GI:52595784
KEYWORDS EST.
SOURCE Petunia x hybrida
ORGANISM Petunia x hybrida
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Petunia.
1 (bases 1 to 495)
Tel: 352-392-1831 x370
Fax: 352-392-3870
Email: dclark@mail.iffas.ufl.edu
Contact Dr. Clark (dclark@mail.iffas.ufl.edu) for clone information
Seq primer: T3 primer.

REFERENCE AUTHORS
Clark, D., Underwood, B., Loucas, H., Farmerie, W., Jones, M. and
Petunia x hybrida EST collection

TITLE Unpublished (2004)
JOURNAL Contact: David Clark
COMMENT UF Horticulture Biotechnology Lab
University of Florida
Environmental Horticulture Department, 1545 Fifield Hall, Box
110670, Gainesville, FL 32611-0670, USA
Tel: 352-392-1831 x370
Fax: 352-392-3870
Email: dclark@mail.iffas.ufl.edu
Contact Dr. Clark (dclark@mail.iffas.ufl.edu) for clone information
Seq primer: T3 primer.

FEATURES

source

1..495
Location/Qualifiers
/organism="Petunia x hybrida"
/mol_type="mRNA"
/cultivar="Mitchell Diploid (aka. Mitchell, aka W15 in Europe)"
/db_xref="taxon:4102"
/clone="Petunia-resq1Teet-D12"
/tissue_type="all floral organs"
/lab_host="lambda ZAPII unidirectional"
/clone_lib="petunia floral development cDNA library"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; supplier: Petunia x hybrida cv. Mitchell Diploid plants were grown from seeds to a fully flowering stage under standard greenhouse conditions. Ten entire flowers of six developmental stages were collected on the same day from plants grown in standard greenhouses. The flower stages were as follows in chronological order from youngest to oldest: stage 1 - no color in corolla; corolla 0.5 inches long stage 2 - first sign of color in corolla; corolla .75-1 inches long stage 3 - fully elongated corolla (not open); corolla 1.5 inches long stage 4 - fully open corolla; anthers not yet dehiscent stage 5 - fully open corolla; freshly anthesed, bright yellow

LOCUS CV294106 555 bp mRNA linear EST 23-SEP-2004
DEFINITION EST862483 petunia floral post-ethylene cDNA library Petunia x
hybrida cDNA clone Petunia-C2H4-16-G01 5' end, mRNA sequence.
ACCESSION CV294106
VERSION CV294106.1 GI:52583065
KEYWORDS EST.
SOURCE Petunia x hybrida
ORGANISM Petunia x hybrida
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Petunia.
1 (bases 1 to 555)
Shibuya,K., Underwood,B., Loucas,H., Farmerie,W., Jones,M. and
Clark,D.
REFERENCE
AUTHORS
TITLE Petunia x hybrida EST collection
JOURNAL Unpublished (2004)
COMMENT Contact: David Clark
UF Horticulture Biotechnology Lab
University of Florida
Environmental Horticulture Department, 1545 Fifield Hall, Box
110670, Gainesville, FL 32611-0670, USA
Tel: 352-392-1831 x370
Fax: 352-392-3870
Email: dclark@mail.ifas.ufl.edu
Contact Dr. Clark (dclark@mail.ifas.ufl.edu) for clone information
Seg primer: T3 primer.
Location/Qualifiers
FEATURES
source
1..555
/organism="Petunia x hybrida"
/mol_type="mRNA"
/cultivar="Mitchell Diploid (aka. Mitchell, aka W15 in
Europe)"
/db_xref="taxon:4102"
/clone="Petunia-C2H4-16-G01"
/tissue_type="all floral organs"
/lab_host="lambda ZAPII unidirectional"
/clone_lib="Petunia floral development cDNA library"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; supplier: Petunia x hybrida cv. Mitchell Diploid
plants were grown from seeds to a fully flowering stage
under standard greenhouse conditions. Flowers at anthesis
stage were excised from plants grown in standard
greenhouses and treated with 2ppm exogenous ethylene.
Entire flowers were collected at 0, 5, 10, 15, 24, and 36
hours after treatment. Total RNA was extracted from each
sample, and 100 micrograms of each sample was combined for
subsequent poly A+ mRNA selection and cDNA synthesis."

ORIGIN

Query Match 54.5%; Score 25.6; DB 7; Length 555;
Best Local Similarity 73.8%; Pred. No. 6.1e+02;
Matches 31; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 6 GGCTGATTAGAAATAAATTGATCACCATTTCAAATTT 47
|||||
Db 264 GGCTGCATCACAAAATTAGTTGATCACCCTTCACAAATGTT 223
|||||

RESULT 22
LOCUS CV297810 614 bp mRNA linear EST 23-SEP-2004
DEFINITION EST866187 petunia floral development cDNA library Petunia x hybrida
cDNA clone Petunia-DeVA-23-A10 5' end, mRNA sequence.
ACCESSION CV297810
VERSION CV297810.1 GI:52590467
KEYWORDS EST.
SOURCE Petunia x hybrida
ORGANISM Petunia x hybrida
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Petunia.
1 (bases 1 to 614)
Shibuya,K., Underwood,B., Loucas,H., Farmerie,W., Jones,M. and

TITLE Petunia x hybrida EST collection
JOURNAL Unpublished (2004)
COMMENT Contact: David Clark
UF Horticulture Biotechnology Lab
University of Florida
Environmental Horticulture Department, 1545 Fifield Hall, Box
110670, Gainesville, FL 32611-0670, USA
Tel: 352-392-1831 x370
Fax: 352-392-3870
Email: dclark@mail.ifas.ufl.edu
Contact Dr. Clark (dclark@mail.ifas.ufl.edu) for clone information
Seg primer: T3 primer.
Location/Qualifiers
FEATURES
source
1..614
/organism="Petunia x hybrida"
/mol_type="mRNA"
/cultivar="Mitchell Diploid (aka. Mitchell, aka W15 in
Europe)"
/db_xref="taxon:4102"
/clone="Petunia-DeVA-23-A10"
/tissue_type="all floral organs"
/lab_host="lambda ZAPII unidirectional"
/clone_lib="Petunia floral development cDNA library"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; supplier: Petunia x hybrida cv. Mitchell Diploid
plants were grown from seeds to a fully flowering stage
under standard greenhouse conditions. Ten entire flowers
of six developmental stages were collected on the same day
from plants grown in standard greenhouses. The flower
stages were as follows in chronological order from
youngest to oldest: stage 1 - no color in corolla; corolla
0.5 inches long stage 2 - first sign of color in corolla;
corolla .75-1 inches long stage 3 - fully elongated
corolla (not open); corolla 1.5 inches long stage 4 -
fully open corolla; anthers not yet dehiscent stage 5 -
fully open corolla; freshly anthesis, bright yellow
pollen; wet stigma stage 6 - pre-senescent; yellowing of
corolla tube; dry brown pollen (if present); stigma dry.
Total RNA was extracted from each sample, and 100
micrograms of each sample was combined for subsequent poly
A+ mRNA selection and cDNA synthesis."

ORIGIN

Query Match 54.5%; Score 25.6; DB 7; Length 614;
Best Local Similarity 73.8%; Pred. No. 6.1e+02;
Matches 31; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 6 GGCTGATTAGAAATAAATTGATCACCATTTCAAATTT 47
|||||
Db 250 GGCTGCATCACAAAATTAGTTGATCACCCTTCACAAATGTT 209
|||||

RESULT 23
LOCUS CV293638 620 bp mRNA linear EST 23-SEP-2004
DEFINITION EST862015 petunia floral post-ethylene cDNA library Petunia x
hybrida cDNA clone Petunia-C2H4-8RR-D11 5' end, mRNA sequence.
ACCESSION CV293638
VERSION CV293638.1 GI:52582127
KEYWORDS EST.
SOURCE Petunia x hybrida
ORGANISM Petunia x hybrida
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Petunia.
1 (bases 1 to 620)
Shibuya,K., Underwood,B., Loucas,H., Farmerie,W., Jones,M. and
Clark,D.
REFERENCE
AUTHORS
TITLE Petunia x hybrida EST collection
JOURNAL Unpublished (2004)
COMMENT Contact: David Clark
UF Horticulture Biotechnology Lab

University of Florida
Environmental Horticulture Department, 1545 Fifield Hall, Box
110670, Gainesville, FL 32611-0670, USA
Tel: 352-392-1831 x370
Fax: 352-392-3870
Email: dclark@mail.ufl.edu
Contact Dr. Clark (dclark@mail.ufl.edu) for clone information
Seq primer: T3 primer.
Location/Qualifiers

FEATURES

source
1..620
/organism="Petunia x hybrida"
/mol_type="mRNA"
/cultivar="Mitchell Diploid (aka. Mitchell, aka W115 in Europe)"
/db_xref="taxon:4102"
/tissue_type="all floral organs"
/lab_host="lambda ZAP11 unidirectional"
/clone_lib="petunia floral post-ethylene cDNA library"
/note="vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; supplier: Petunia x hybrida cv. Mitchell Diploid plants were grown from seeds to a fully flowering stage under standard greenhouse conditions. Flowers at anthesis stage were excised from plants grown in standard greenhouses and treated with 2ppm exogenous ethylene. Entire flowers were collected at 0, 5, 10, 15, 24, and 36 hours after treatment. Total RNA was extracted from each sample, and 100 micrograms of each sample was combined for subsequent poly A+ mRNA selection and cDNA synthesis."

ORIGIN

Query Match 54.5%; Score 25.6; DB 7; Length 620;
Best Local Similarity 73.8%; Pred. No. 6.1e+02;
Matches 31; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

6 GCCTGATTGAGAAATTAATTTGATCACCATTTCGAATTTT 47
|||||
243 GGCTGCATCACAAATTAATTTGATCACCCTTCACAAATGTT 202

RESULT 24
CV294470/c 630 bp mRNA linear EST 23-SEP-2004
LOCUS
DEFINITION EST82847 petunia floral post-ethylene cDNA library Petunia x hybrida cDNA clone Petunia-C2H4-21-H04 5' end, mRNA sequence.
ACCESSION CV294470
VERSION CV294470.1 GI:52583793
KEYWORDS
SOURCE
ORGANISM
Petunia x hybrida
Petunia x hybrida
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Petunia.
1 (bases 1 to 630)
Shibuya, K., Underwood, B., Loucas, H., Farmerie, W., Jones, M. and Clark, D.
Petunia x hybrida EST collection
Unpublished (2004)
Contact: David Clark
UF Horticulture Biotechnology Lab
University of Florida
Environmental Horticulture Department, 1545 Fifield Hall, Box 110670, Gainesville, FL 32611-0670, USA
Tel: 352-392-1831 x370
Fax: 352-392-3870
Email: dclark@mail.ufl.edu
Contact Dr. Clark (dclark@mail.ufl.edu) for clone information
Seq primer: T3 primer.
Location/Qualifiers

TITLE
JOURNAL
COMMENT

REFERENCE
AUTHORS

FEATURES

source
1..630
/organism="Petunia x hybrida"
/mol_type="mRNA"
/cultivar="Mitchell Diploid (aka. Mitchell, aka W115 in

Europe)"
/db_xref="taxon:4102"
/clone="Petunia-C2H4-21-H04"
/tissue_type="all floral organs"
/lab_host="lambda ZAP11 unidirectional"
/clone_lib="petunia floral post-ethylene cDNA library"
/note="vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; supplier: Petunia x hybrida cv. Mitchell Diploid plants were grown from seeds to a fully flowering stage under standard greenhouse conditions. Flowers at anthesis stage were excised from plants grown in standard greenhouses and treated with 2ppm exogenous ethylene. Entire flowers were collected at 0, 5, 10, 15, 24, and 36 hours after treatment. Total RNA was extracted from each sample, and 100 micrograms of each sample was combined for subsequent poly A+ mRNA selection and cDNA synthesis."

ORIGIN

Query Match 54.5%; Score 25.6; DB 7; Length 630;
Best Local Similarity 73.8%; Pred. No. 6.1e+02;
Matches 31; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

6 GCCTGATTGAGAAATTAATTTGATCACCATTTCGAATTTT 47
|||||
327 GGCTGCATCACAAATTAATTTGATCACCCTTCACAAATGTT 286

RESULT 25
CV293898/c 638 bp mRNA linear EST 23-SEP-2004
LOCUS
DEFINITION EST828275 petunia floral post-ethylene cDNA library Petunia x hybrida cDNA clone Petunia-C2H4-15-D09 5' end, mRNA sequence.
ACCESSION CV293898
VERSION CV293898.1 GI:52582649
KEYWORDS
SOURCE
ORGANISM
Petunia x hybrida
Petunia x hybrida
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Petunia.
1 (bases 1 to 638)
Shibuya, K., Underwood, B., Loucas, H., Farmerie, W., Jones, M. and Clark, D.
Petunia x hybrida EST collection
Unpublished (2004)
Contact: David Clark
UF Horticulture Biotechnology Lab
University of Florida
Environmental Horticulture Department, 1545 Fifield Hall, Box 110670, Gainesville, FL 32611-0670, USA
Tel: 352-392-1831 x370
Fax: 352-392-3870
Email: dclark@mail.ufl.edu
Contact Dr. Clark (dclark@mail.ufl.edu) for clone information
Seq primer: T3 primer.
Location/Qualifiers

FEATURES

source
1..638
/organism="Petunia x hybrida"
/mol_type="mRNA"
/cultivar="Mitchell Diploid (aka. Mitchell, aka W115 in Europe)"
/db_xref="taxon:4102"
/clone="Petunia-C2H4-15-D09"
/tissue_type="all floral organs"
/lab_host="lambda ZAP11 unidirectional"
/clone_lib="petunia floral post-ethylene cDNA library"
/note="vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; supplier: Petunia x hybrida cv. Mitchell Diploid plants were grown from seeds to a fully flowering stage under standard greenhouse conditions. Flowers at anthesis stage were excised from plants grown in standard greenhouses and treated with 2ppm exogenous ethylene. Entire flowers were collected at 0, 5, 10, 15, 24, and 36

hours after treatment. Total RNA was extracted from each sample, and 100 micrograms of each sample was combined for subsequent poly A+ mRNA selection and cDNA synthesis."

ORIGIN

Query Match 54.5%; Score 25.6; DB 7; Length 638;
 Best Local Similarity 73.8%; Pred. No. 6.1e+02;
 Matches 31; Conservative 1; Mismatches 10; Indels 0; Gaps 0;
 QY 6 GGCTTGATTAGAAATTAATTAATTGATCACCATTTCGAATTT 47
 ||||| ||||| ||||| :||||||| || ||||| ||
 Db 629 GGCTCGATCACAATAATTAGTTGATCACCCTTCACAATGTT 588

Search completed: June 2, 2005, 09:49:49
 Job time : 3145 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2005, 08:30:33 ; Search time 126 Seconds
(without alignments)
610.357 Million cell updates/sec

Title: US-10-349-143-1

Perfect score: 47

Sequence: 1 ccacgcgtcgtatgaat.....gatcacatttcaatttt 47

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/6C_COMB.seq:*
6: /cgn2_6/prodata/1/ina/6D_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46.6	99.1	47	US-09-422-978-1	Sequence 1, Appl1
2	25.2	53.6	601	US-09-949-016-52779	Sequence 52779, A
3	25.2	53.6	601	US-09-949-016-52780	Sequence 52780, A
4	25.2	53.6	187580	US-09-949-016-13366	Sequence 13366, A
5	24.6	52.3	37195	US-09-949-016-13364	Sequence 13264, A
6	24.2	51.5	273	US-09-248-796A-7359	Sequence 7359, Ap
7	24.2	51.5	993	US-09-328-352-3876	Sequence 3876, Ap
8	24	51.1	96922	US-09-949-016-17061	Sequence 17061, A
9	23.6	50.2	601	US-09-949-016-52894	Sequence 52894, A
10	23.6	50.2	90866	US-09-949-016-13271	Sequence 13271, A
11	23.2	49.4	601	US-09-949-016-185469	Sequence 185469, A
12	23.2	49.4	601	US-09-949-016-185470	Sequence 185470, A
13	23.2	49.4	1978	US-09-949-016-1358	Sequence 1358, Ap
14	23.2	49.4	2064	US-09-248-796A-1897	Sequence 1897, Ap
15	23.2	49.4	24538	US-09-949-016-13100	Sequence 13100, A
16	23.2	49.4	44971	US-09-949-016-17049	Sequence 17049, A
17	23	48.9	76164	US-09-949-016-12288	Sequence 12288, A
18	23	48.9	76165	US-09-949-016-12288	Sequence 12288, A
19	22.8	48.5	31491	US-09-949-016-14005	Sequence 14005, A
20	22.8	48.5	31491	US-09-949-016-14005	Sequence 14005, A
21	22.6	48.1	459	US-08-856-171E-1138	Sequence 1138, Ap
22	22.6	48.1	459	US-08-856-171E-1138	Sequence 1138, Ap
23	22.6	48.1	459	US-08-856-171E-1138	Sequence 1138, Ap
24	22.6	48.1	1209	US-09-949-016-14510	Sequence 14510, A
25	22.6	48.1	57726	US-08-336-165A-55	Sequence 55, Appl1
26	22.6	48.1	61178	US-09-949-016-16430	Sequence 16430, A
27	22.6	48.1	181429	US-09-949-016-17369	Sequence 17369, A

28	22.6	48.1	181430	US-09-949-016-15772	Sequence 15772, A
29	22.6	48.1	205163	US-09-949-016-17009	Sequence 17009, A
30	22.6	48.1	236964	US-09-949-016-15753	Sequence 15753, A
31	22.4	47.7	348	US-09-583-110-1030	Sequence 1030, Ap
32	22.4	47.7	492	US-09-107-433-1285	Sequence 1285, Ap
33	22.4	47.7	745	US-09-495-050A-183	Sequence 183, App
34	22.4	47.7	3257	US-09-949-016-1496	Sequence 1496, Ap
35	22.4	47.7	3986	US-09-949-016-24	Sequence 24, Appl1
36	22.4	47.7	46603	US-09-949-016-15636	Sequence 15636, A
37	22.4	47.7	63558	US-09-949-016-13238	Sequence 13238, A
38	22.4	47.7	64489	US-09-949-016-11766	Sequence 11766, A
39	22.2	47.2	601	US-09-949-016-30145	Sequence 30145, A
40	22.2	47.2	601	US-09-949-016-37382	Sequence 37382, A
41	22.2	47.2	601	US-09-949-016-65350	Sequence 65350, A
42	22.2	47.2	601	US-09-949-016-142405	Sequence 142405, A
43	22.2	47.2	601	US-09-949-016-142595	Sequence 142595, A
44	22.2	47.2	981	US-09-671-317-38	Sequence 38, Appl1
45	22.2	47.2	3001	US-09-539-333D-127	Sequence 127, App
46	22.2	47.2	28213	US-09-949-016-12738	Sequence 12738, A
47	22.2	47.2	28216	US-09-949-016-13652	Sequence 13652, A
48	22.2	47.2	55886	US-09-949-016-15129	Sequence 15129, A
49	22.2	47.2	62311	US-09-949-016-14582	Sequence 14582, A
50	22.2	47.2	95890	US-09-949-016-16412	Sequence 16412, A
51	22.2	47.2	211049	US-09-949-016-15770	Sequence 15770, A
52	22	46.8	571	US-09-854-133-307	Sequence 307, App
53	22	46.8	601	US-09-949-016-124854	Sequence 124854, A
54	22	46.8	601	US-09-949-016-136288	Sequence 136288, A
55	22	46.8	601	US-09-949-016-151005	Sequence 151005, A
56	22	46.8	49416	US-09-949-016-15234	Sequence 15234, A
57	22	46.8	50950	US-09-949-016-16559	Sequence 16559, A
58	22	46.8	91772	US-09-949-016-15568	Sequence 15568, A
59	22	46.8	132438	US-09-949-016-14349	Sequence 14349, A
60	22	46.8	132438	US-09-949-016-14350	Sequence 14350, A
61	22	46.8	151089	US-09-949-016-14348	Sequence 14348, A
62	22	46.8	192302	US-09-949-016-15270	Sequence 15270, A
63	22	46.8	239527	US-09-949-016-15980	Sequence 15980, A
64	22	46.8	524032	US-09-949-016-16928	Sequence 16928, A
65	22	46.8	524032	US-09-949-016-16930	Sequence 16930, A
66	22	46.8	524032	US-09-949-016-16932	Sequence 16932, A
67	22	46.8	524032	US-09-949-016-16931	Sequence 16931, A
68	22	46.8	529885	US-09-949-016-14340	Sequence 14340, A
69	22	46.8	529885	US-09-949-016-14341	Sequence 14341, A
70	22	46.8	529885	US-09-949-016-14342	Sequence 14342, A
71	22	46.8	529885	US-09-949-016-14343	Sequence 14343, A
72	22	46.8	529885	US-09-949-016-14344	Sequence 14344, A
73	22	46.8	529885	US-09-949-016-14345	Sequence 14345, A
74	22	46.8	529885	US-09-949-016-14346	Sequence 14346, A
75	22	46.8	529885	US-09-949-016-14347	Sequence 14347, A
76	21.8	46.4	472	US-08-956-171E-948	Sequence 948, App
77	21.8	46.4	472	US-08-956-171E-948	Sequence 948, App
78	21.8	46.4	601	US-09-949-001-140	Sequence 140, App
79	21.8	46.4	601	US-09-949-001-1606	Sequence 606, App
80	21.8	46.4	861	US-09-276-531-90	Sequence 90, Appl1
81	21.8	46.4	1368	US-09-540-236-731	Sequence 731, Appl
82	21.8	46.4	5208	US-08-781-891-71	Sequence 71, Appl1
83	21.8	46.4	5208	US-09-791-211-11	Sequence 11, Appl1
84	21.8	46.4	5208	US-09-618-166-70	Sequence 70, Appl1
85	21.8	46.4	5308	US-09-949-001-9	Sequence 9, Appl1
86	21.8	46.4	5308	US-09-949-001-13	Sequence 13, Appl1
87	21.8	46.4	5308	US-09-949-001-15	Sequence 15, Appl1
88	21.8	46.4	7888	US-09-573-080A-193	Sequence 193, Appl1
89	21.8	46.4	10660	US-08-567-803B-8	Sequence 8, Appl1
90	21.8	46.4	10660	US-09-041-886-16	Sequence 16, Appl1
91	21.8	46.4	18367	US-09-949-016-14065	Sequence 14065, A
92	21.8	46.4	30371	US-09-949-016-15595	Sequence 15395, A
93	21.8	46.4	48328	US-09-949-016-16935	Sequence 16935, A
94	21.8	46.4	48328	US-09-596-002-27	Sequence 27, Appl1
95	21.8	46.4	87350	US-08-781-891-79	Sequence 79, Appl1
96	21.8	46.4	87350	US-09-618-166-79	Sequence 79, Appl1
97	21.8	46.4	87343	US-09-791-211-3	Sequence 3, Appl1
98	21.8	46.4	106418	US-09-949-016-13974	Sequence 13974, A
99	21.8	46.4	111335	US-09-949-016-15282	Sequence 15382, A
100	21.8	46.4	132871	US-09-949-016-13863	Sequence 13863, A

ALIGNMENTS

RESULT 1
US-09-422-978-1; Sequence 1, Application US/09422978
; Patent No. 6537751

; GENERAL INFORMATION:

; APPLICANT: Cohen, Daniel

; APPLICANT: Blumenfeld, Marta

; APPLICANT: Chumakov, Ilya

; TITLE OF INVENTION: Biallelic markers for use in constructing a high density....

; FILE REFERENCE: GENSET.020CPI

; CURRENT APPLICATION NUMBER: US/09/422,978

; PRIOR FILING DATE: 1999-10-20

; EARLIER APPLICATION NUMBER: US 09/298,850

; EARLIER FILING DATE: 1999-04-21

; EARLIER APPLICATION NUMBER: US 60/109,732

; EARLIER FILING DATE: 1998-11-23

; EARLIER APPLICATION NUMBER: US 60/082,614

; NUMBER OF SEQ ID NOS: 11796

; SEQ ID NO 1

; LENGTH: 47

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE:

; NAME/KEY: allele

; LOCATION: 24

; OTHER INFORMATION: 99-109-224 : polymorphic base G or C

US-09-422-978-1

Query Match 99.1%; Score 46.6; DB 4; Length 47;

Best Local Similarity 100.0%; Pred. No. 2.4e-07;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCACAGGCTTGATTAGAAATAAATTGATCACCATTTCAAATTTT 47

Db 1 CCACAGGCTTGATTAGAAATAAATTGATCACCATTTCAAATTTT 47

RESULT 2
US-09-949-016-52779

; Sequence 52779, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: C1001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 52779

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-52779

Query Match 53.6%; Score 25.2; DB 4; Length 601;

Best Local Similarity 71.7%; Pred. No. 11;

Matches 33; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Db 296 CACCAATGTTTAAATAATATAGTTCATGCCCATTTCAATTTCT 341

RESULT 3
US-09-949-016-52780

; Sequence 52780, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: C1001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 52780

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-52780

Query Match 53.6%; Score 25.2; DB 4; Length 601;

Best Local Similarity 71.7%; Pred. No. 11;

Matches 33; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 2 CACAGGCTTGATTAGAAATAAATTGATCACCATTTCAAATTTT 47

Db 434 CACCAATGTTTAAATAATATAGTTCATGCCCATTTCAATTTCT 479

RESULT 4
US-09-949-016-13266/C

; Sequence 13266, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: C1001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 13266

; LENGTH: 187580

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(187580)

; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-13266

Query Match 53.6%; Score 25.2; DB 4; Length 187580;

Best Local Similarity 71.7%; Pred. No. 24;

Matches 33; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 2 CACAGGCTTGATTAGAAATAAATTGATCACCATTTCAAATTTT 47

Db 2 CACAGGCTTGATTAGAAATAAATTGATCACCATTTCAAATTTT 47

Db 178619 CACCATCATGTTTAAATATATAGTTCAGCCCATTTTCAATTCT 178574

RESULT 5

US-09-949-016-13264/c
; Sequence 13264, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13264
; LENGTH: 37195
; TYPE: DNA
; ORGANISM: Human
; FEATURE: NAME/KEY: misc_feature
; LOCATION: (1)...(37195)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13264

Query Match 52.3%; Score 24.6; DB 4; Length 37195;
Best Local Similarity 70.2%; Pred. No. 31;
Matches 33; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 CCACAGGCTTGATAGAAATTAATTTGATCACCATTTCGAATTTT 47
Db 10945 CCAGAGGCTTGATGCTGGGAAAGAAATATGATTACCAATTTTGAATTTT 10899

RESULT 6

US-09-248-796A-7359/c
; Sequence 7359, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 7359
; LENGTH: 273
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-7359

Query Match 51.5%; Score 24.2; DB 4; Length 273;
Best Local Similarity 68.1%; Pred. No. 21;
Matches 32; Conservative 1; Mismatches 14; Indels 0; Gaps 0;

Qy 1 CCACAGGCTTGATAGAAATTAATTTGATCACCATTTCGAATTTT 47
Db 211 CCACAGTTTATTGGAAATTAATGCGATCCCAACTTCAATCTT 165

RESULT 7

US-09-328-352-3876/c
; Sequence 3876, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 3876
; LENGTH: 993
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-3876

Query Match 51.5%; Score 24.2; DB 4; Length 993;
Best Local Similarity 68.1%; Pred. No. 25;
Matches 32; Conservative 1; Mismatches 14; Indels 0; Gaps 0;

Qy 1 CCACAGGCTTGATAGAAATTAATTTGATCACCATTTCGAATTTT 47
Db 205 CACAGGTTTACTGAGAGATTAATGCTGATTTACTATTGAGACTTT 159

RESULT 8

US-09-949-016-17061
; Sequence 17061, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17061
; LENGTH: 96922
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17061

Query Match 51.1%; Score 24; DB 4; Length 96922;
Best Local Similarity 75.0%; Pred. No. 58;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 8 CTTGATTAGAAATTAATTTGATCACCATTTCGAATTTT 47
Db 34578 CATTAATTAGAAATTAATTAATTTGATCCGATATTTTAAATTTT 34617

RESULT 9

US-09-949-016-52894
; Sequence 52894, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20

;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 52894
;; LENGTH: 601
;; TYPE: DNA
;; ORGANISM: Human
US-09-949-016-52894

Query Match 50.2%; Score 23.6; DB 4; Length 601;
Best Local Similarity 76.3%; Pred. No. 38;
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 9 TTGATTAGAAATAAATTGATCACCATTTCATAATT 46
Db 204 TTGATTATAAATAAATTTCATATCAATTTCTATT 241

RESULT 10
US-09-949-016-13271/C
; Sequence 13271, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 13271
; LENGTH: 90876
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(90876)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13271

Query Match 50.2%; Score 23.6; DB 4; Length 90876;
Best Local Similarity 76.3%; Pred. No. 79;
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 9 TTGATTAGAAATAAATTGATCACCATTTCATAATT 46
Db 24298 TTGATTATAAATAAATTTCATATCAATTTCTATT 24261

RESULT 11
US-09-949-016-185469
; Sequence 185469, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 185469
;; LENGTH: 601
;; TYPE: DNA
;; ORGANISM: Human
US-09-949-016-185469

Query Match 49.4%; Score 23.2; DB 4; Length 601;
Best Local Similarity 73.7%; Pred. No. 53;
Matches 28; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

Qy 5 AGGCTTGATTAGAAATAAATTGATCACCATTTCAA 42
Db 148 AGGATTCAGTAGAATAAAGTTTATTTGCTTTCGA 185

RESULT 12
US-09-949-016-185470
; Sequence 185470, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 185470
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-185470

Query Match 49.4%; Score 23.2; DB 4; Length 601;
Best Local Similarity 73.7%; Pred. No. 53;
Matches 28; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

Qy 5 AGGCTTGATTAGAAATAAATTGATCACCATTTCAA 42
Db 140 AGGATTCAGTAGAATAAAGTTTATTTGCTTTCGA 177

RESULT 13
US-09-949-016-1358
; Sequence 1358, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1358
LENGTH: 1978
TYPE: DNA
ORGANISM: Human
US-09-949-016-1358

Query Match 49.4%; Score 23.2; DB 4; Length 1978;
Best Local Similarity 83.3%; Pred. No. 63;
Matches 25; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 10 TGATTGAAATTAATTTGATCACCATT 39
DB 1939 TGACTTGAATTAACCTTGACACCAATTT 1968

RESULT 14
US-09-248-796A-1897
Sequence 1897, Application US/09248796A
Patent No. 6747137

GENERAL INFORMATION:
APPLICANT: Keith Weinstein et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12

PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/074,725

PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/096,409

NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 1897

LENGTH: 2064

TYPE: DNA
ORGANISM: Candida albicans
US-09-248-796A-1897

Query Match 49.4%; Score 23.2; DB 4; Length 2064;
Best Local Similarity 73.7%; Pred. No. 63;
Matches 28; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 9 TTGATTGAAATTAATTTGATCACCATTTCGAATTT 46
DB 1609 TTGATTGACATGAACTTTATCTACATTTAAACTT 1646

RESULT 15
US-09-949-016-13100
Sequence 13100, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13100

LENGTH: 24538
TYPE: DNA

ORGANISM: Human
US-09-949-016-13100

Query Match 49.4%; Score 23.2; DB 4; Length 24538;
Best Local Similarity 83.3%; Pred. No. 90;

Matches 25; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 10 TGATTGAAATTAATTTGATCACCATT 39
DB 22499 TGACTTGAATTAACCTTGACACCAATTT 22528

RESULT 16
US-09-949-016-17049
Sequence 17049, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498

NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 17049
LENGTH: 44971

TYPE: DNA
ORGANISM: Human

FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...((44971))

OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17049

Query Match 49.4%; Score 23.2; DB 4; Length 44971;
Best Local Similarity 73.7%; Pred. No. 98;
Matches 28; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 5 AGCTTGATTGAAATTAATTTGATCACCATTTCGA 42
DB 31128 AGGATTGAGTGAATTAAGTTTATTCGTTGCA 31165

RESULT 17
US-09-949-016-12288/c
Sequence 12288, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12288

LENGTH: 76164
TYPE: DNA

ORGANISM: Human
US-09-949-016-12288

Query Match 48.9%; Score 23; DB 4; Length 76164;
Best Local Similarity 74.4%; Pred. No. 1.2e+02;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

Oy      9  TTGATTAGAAATTAASTTTTGATCACCATTTCAAATTTT  47
          || ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    40627  TTCAATTAAATTTAAATTAATTACATTTTAAAACTTT  40589

```

```

1      RESULT 18
2      US-09-949-016-14005/c
3      ; Sequence 14005, Application US/09949016
4      ; Patent No. 6812339
5      ; GENERAL INFORMATION:
6      ; APPLICANT: VENTER, J. Craig et al.
7      ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
8      ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
9      ; FILE REFERENCE: CLO01307
10     ; CURRENT APPLICATION NUMBER: US/09/949, 016
11     ; CURRENT FILING DATE: 2000-04-14
12     ; PRIOR APPLICATION NUMBER: 60/241,755
13     ; PRIOR FILING DATE: 2000-10-20
14     ; PRIOR APPLICATION NUMBER: 60/237,768
15     ; PRIOR FILING DATE: 2000-10-03
16     ; PRIOR APPLICATION NUMBER: 60/231,498
17     ; PRIOR FILING DATE: 2000-09-08
18     ; NUMBER OF SEQ ID NOS: 207012
19     ; SOFTWARE: FastSeq for Windows Version 4.0
20     ; SEQ ID NO 14005
21     ; LENGTH: 76165
22     ; TYPE: DNA
23     ; ORGANISM: Human
24     ; US-09-949-016-14005

```

Qy	9	TTGATAGAAATAAATTGGATGCCATTTTCAAAATTTT	47
Db	40627	TTCAATTAAATAATTAATAATTAACATTTTAAACCTTT	40589
		Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0	
		Query Match 94.9%; Score 23; DB 4; Length 76165; Best Local Similarity 74.4%; Pred. No. 1.2e-02;	

```

RESULT 19
US-09-360-186-1/c
: Sequence 1, Application US/09360186
: Patent No. 6262343
: GENERAL INFORMATION:
: APPLICANT: Staslawicz, et al.
: TITLE OF INVENTION: B62 Resistance Gene
: FILE REFERENCE: 50687
: CURRENT APPLICATION NUMBER: US/09/360,186
: CURRENT FILING DATE: 1999-07-23
: EARLIER APPLICATION NUMBER: 60/093,957
: EARLIER FILING DATE: 1998-07-23
: NUMBER OF SEQ ID NOS: 9
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 31491
: TYPE: DNA
: ORGANISM: Capsicum annuum
: US-09-360-186-1

```

Query Match	48.5%	Score 22.8	DB 3	length 31491
Best Local Similarity	71.4%	Pred. No. 1.3e+02		
Matches 30; Conservative	0	Mismatches 12	Indels 0	Gaps 0

RESULT 20
US-09-864-680A-1/c
; Sequence 1, Application US/09864680A
; Patent No. 6762285

```

1  GENERAL INFORMATION:
2  APPLICANT: Staskawicz, Brian J
3  APPLICANT: Dahlbeck, Douglas
4  APPLICANT: Tai, Thomas H
5  TITLE OF INVENTION: B82 RESISTANCE GENE
6  FILE REFERENCE: 42250/234021 (5830-44)
7  CURRENT APPLICATION NUMBER: US/09/864,680A
8  CURRENT FILING DATE: 2001-05-24
9  PRIOR APPLICATION NUMBER: US 09/360,186
10 PRIOR FILING DATE: 1999-07-23
11 PRIOR APPLICATION NUMBER: US 60/093,957
12 PRIOR FILING DATE: 1998-07-23
13 NUMBER OF SEQ ID NOS: 14
14 SOFTWARE: PatentIn version 3.1
15 SEQ ID NO 1
16 LENGTH: 31491
17 TYPE: DNA
18 ORGANISM: Capsicum annuum
19 US-09-864-680A-1

```

Query	3	ACAGGCTTGATTAGCAATTAASSTTGATCCACCATTTTCAAT	44
db	27659	AGACCATGCAATTGAAATAAATATATGTCACCATCTTGAAAT	27618

Query Match 30; Conservative 0; Mismatches 12; Indels 0;
 48.5%; Score 22.8; DB 4; length 31491;
 Best Local Similarity 71.4%; Pred. No. 1.3e+02;
 Matches 30; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

RESULT 21
 US-08-956-171E-1138
 Sequence 1138, Application US/08956171E
 Patent No. 6593114
 GENERAL INFORMATION:
 APPLICANT: Charles Kunsch
 Gil H. Choi
 Patrick S. Dillon
 Craig A. Rosen
 Steven C. Barash
 Michael R. Fannon
 TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
 NUMBER OF SEQUENCES: 5256
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville
 STATE: Maryland
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 COMPUTER: HP Vectra 486/33
 OPERATING SYSTEM: MSDOS version 6.2
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/956,171E
 FILING DATE: 20-Oct-1997
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/009,861
 FILING DATE: January 5, 1996
 APPLICATION NUMBER: 08/781,986
 FILING DATE: January 3, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Mark J. Hyman
 REGISTRATION NUMBER: 46,789
 REFERENCE/DOCKET NUMBER: PB248P1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (240) 314-1224
 TELEFAX: (301) 309-8439
 INFORMATION FOR SEQ ID NO: 1138:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 459 base pairs

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1138:
US-08-956-171E-1138

Query Match 48.1%; Score 22.6; DB 4; Length 459;
Best Local Similarity 68.9%; Pred. No. 82;
Matches 31; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 3 ACAGCGTTGATTGAAATAAATTGATCACCATTTCCTCAATTTT 47
DB 317 AAAAAGTTTGTCACAAATAAATCTATCTCAATTTTAAATAT 361

RESULT 22
US-08-781-986A-1138
Sequence 1138, Application US/08781986A
Patent No. 6737248

GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Benson, Bob

REGISTRATION NUMBER: 30,446

REFERENCE/DOCKET NUMBER: PB248PP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 1138:

SEQUENCE CHARACTERISTICS:

LENGTH: 459 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-781-986A-1138

Query Match 48.1%; Score 22.6; DB 4; Length 459;
Best Local Similarity 68.9%; Pred. No. 82;

Matches 31; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 3 ACAGCGTTGATTGAAATAAATTGATCACCATTTCCTCAATTTT 47
DB 317 AAAAAGTTTGTCACAAATAAATCTATCTCAATTTTAAATAT 361

RESULT 23
US-09-949-016-141510
Sequence 141510, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO: 141510

LENGTH: 601

TYPE: DNA

ORGANISM: Human

US-09-949-016-141510

Query Match 48.1%; Score 22.6; DB 4; Length 601;
Best Local Similarity 71.8%; Pred. No. 85;

Matches 28; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 8 CTGATTGAAATAAATTGATCACCATTTCCTCAATTTT 46
DB 337 CTGATTGAAATAAATTGATCACCATTTCCTCAATTTTAAAT 375

RESULT 24
US-08-936-165A-55/C
Sequence 55, Application US/08936165A
Patent No. 6348582

GENERAL INFORMATION:

APPLICANT: Black, Michael

APPLICANT: Burnham, Martin

APPLICANT: Hodgson, John

APPLICANT: Knowles, David

APPLICANT: Lometto, Michael

APPLICANT: Nicholas, Richard

APPLICANT: Pratt, Julie

APPLICANT: Reichard, Richard

APPLICANT: Rosenberg, Martin

APPLICANT: Ward, Judith

TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,

TITLE OF INVENTION: Polypeptides and Their Uses

NUMBER OF SEQUENCES: 534

CORRESPONDENCE ADDRESS:

ADDRESSEE: SmithKline Beecham Corporation

STREET: 709 Swedeland Road

CITY: King of Prussia

STATE: PA

COUNTRY: USA

ZIP: 19406-0939

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/936,165A

FILING DATE: 24-SEP-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/027,032

FILING DATE: 24-SEP-1996

ATTORNEY/AGENT INFORMATION:

NAME: Glumel, Edward R

REGISTRATION NUMBER: 38,891

REFERENCE/DOCKET NUMBER: P50549

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-270-4478

TELEFAX: 610-270-5090

TELEX:

INFORMATION FOR SEQ ID NO: 55:

/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1209 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: Genomic DNA
US-08-936-165A-55

Query Match 48.1%; Score 22.6; DB 3; Length 1209;
Best Local Similarity 68.9%; Pred. No. 94;
Matches 31; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 3 ACAGGCTTGATTGAAATTAATTAATTTGATCACCATTTCATAATTTT 47
Db 439 AAAAAGCTTTGACAAATTAATTAATTTGATCATTTCATAATTTT 395

RESULT 25

US-09-949-016-16430
; Sequence 16430, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: fastSeq for Windows Version 4.0
; SEQ ID NO 16430
; LENGTH: 57726
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16430

Query Match 48.1%; Score 22.6; DB 4; Length 57726;
Best Local Similarity 68.9%; Pred. No. 1.6e+02;
Matches 31; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 CCACGCGCTTGATTGAAATTAATTAATTTGATCACCATTTCATAATTT 45
Db 48861 CCATTAAGGTAGATTAACATTAATTTTCCGTATTAATTT 48905

Search completed: June 2, 2005, 10:22:43
Job time : 136 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2005, 08:49:54 ; Search time 517 Seconds
(without alignments)
558.857 Million cell updates/sec

Title: US-10-349-143-1

Perfect score: 47
Sequence: 1 ccacagcgtcgttagaatt.....gattcaccatttcaatttt 47

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5706582 seqs, 3073711274 residues

Total number of hits satisfying chosen parameters: 11413164

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 100 summaries

Database : Published Applications NA:*

1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
4: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
5: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
6: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
7: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
8: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq:*
9: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq:*
12: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
17: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
18: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
19: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
20: /cgn2_6/prodata/2/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46.6	99.1	47	US-10-349-143-1	Sequence 1, Appl1
2	46.6	99.1	47	US-10-349-143-1	Sequence 1, Appl1
3	46.6	99.1	47	US-10-349-143-1	Sequence 1, Appl1
4	24.8	52.8	18	US-10-468-582-1	Sequence 1, Appl1
5	24.8	52.8	18	US-10-719-993-29983	Sequence 1, Appl1
6	24.4	51.9	267	US-09-796-692-4373	Sequence 6843, Ap
7	24.4	51.9	267	US-09-796-692-4373	Sequence 29983, Ap
8	24.4	51.9	267	US-10-040-862-4373	Sequence 4373, Ap
9	24.4	51.9	267	US-10-040-862-4373	Sequence 4373, Ap
10	24.4	51.9	267	US-10-057-475B-4373	Sequence 4373, Ap
11	24.4	51.9	267	US-10-057-475B-9383	Sequence 9383, Ap

C 12	24.4	51.9	267	17	US-10-154-884B-4373	Sequence 4373, Ap
C 13	24.4	51.9	267	17	US-10-154-884B-9383	Sequence 9383, Ap
C 14	24.4	51.9	267	17	US-10-764-324-4373	Sequence 4373, Ap
C 15	24.4	51.9	267	18	US-10-764-324-9383	Sequence 9383, Ap
C 16	24.4	51.9	589	13	US-10-027-632-240198	Sequence 240198, A
C 17	24.4	51.9	589	13	US-10-027-632-240199	Sequence 240199, A
C 18	24.4	51.9	589	13	US-10-027-632-240199	Sequence 240199, A
C 19	24.4	51.9	589	13	US-10-027-632-240198	Sequence 240198, A
C 20	24.4	51.9	589	17	US-10-027-632-240199	Sequence 240199, A
C 21	24.4	51.9	589	17	US-10-027-632-240200	Sequence 240200, A
C 22	24.2	51.5	2199	18	US-10-437-963-51548	Sequence 51548, A
C 23	24.2	51.5	5173	15	US-10-311-455-924	Sequence 924, App
C 24	23.8	50.6	728	18	US-10-767-701-2646	Sequence 2646, Ap
C 25	23.8	50.6	1102	17	US-10-425-114-15544	Sequence 15544, A
C 26	23.8	50.6	1105	18	US-10-425-115-83159	Sequence 83159, A
C 27	23.6	50.2	27681	9	US-09-764-869-1997	Sequence 1997, Ap
C 28	23.6	50.2	27681	9	US-09-764-869-1998	Sequence 1998, Ap
C 29	23.6	50.2	27681	14	US-10-091-504-1997	Sequence 1997, Ap
C 30	23.6	50.2	27681	14	US-10-091-504-1998	Sequence 1998, Ap
C 31	23.6	50.2	27681	17	US-10-227-577-1997	Sequence 1997, Ap
C 32	23.6	50.2	27681	17	US-10-227-577-1998	Sequence 1998, Ap
C 33	23.4	49.8	366	18	US-10-425-115-182292	Sequence 182292, A
C 34	23.4	49.8	562	13	US-10-027-632-49686	Sequence 49686, A
C 35	23.4	49.8	562	17	US-10-027-632-49686	Sequence 49686, A
C 36	23.4	49.8	581	13	US-10-027-632-71246	Sequence 71246, A
C 37	23.4	49.8	581	13	US-10-027-632-300279	Sequence 300279, A
C 38	23.4	49.8	581	17	US-10-027-632-300279	Sequence 300279, A
C 39	23.4	49.8	581	17	US-10-027-632-300279	Sequence 300279, A
C 40	23.4	49.8	721	13	US-10-027-632-155815	Sequence 155815, A
C 41	23.4	49.8	721	17	US-10-027-632-155815	Sequence 155815, A
C 42	23.2	49.4	370	17	US-10-424-559-35985	Sequence 35985, A
C 43	23.2	49.4	443	18	US-10-674-124A-8285	Sequence 8285, Ap
C 44	23.2	49.4	546	13	US-10-027-632-278285	Sequence 278285, A
C 45	23.2	49.4	546	17	US-10-027-632-278285	Sequence 278285, A
C 46	23.2	49.4	942	18	US-10-474-792-623	Sequence 623, App
C 47	23.2	49.4	2000	17	US-10-260-239-1809	Sequence 1809, App
C 48	23.2	49.4	2347	17	US-10-108-260A-852	Sequence 852, App
C 49	23.2	49.4	4062	17	US-10-424-559-19022	Sequence 19022, A
C 50	23.2	49.4	34885	19	US-10-741-600-17927	Sequence 17927, A
C 51	23.2	49.4	38295	18	US-10-741-600-5755	Sequence 5755, Ap
C 52	23.2	49.4	43494	13	US-10-741-600-17594	Sequence 17594, A
C 53	23.2	49.4	61197	19	US-10-087-190-1924	Sequence 1924, Ap
C 54	23.2	49.4	90435	18	US-10-322-696-1160	Sequence 160, App
C 55	23.2	49.4	11463	18	US-10-473-397-3	Sequence 3, Appl1
C 56	23.2	49.4	126001	17	US-10-173-492-13	Sequence 13, Appl1
C 57	23	48.9	411	9	US-09-983-965-4059	Sequence 4059, Ap
C 58	23	48.9	683	13	US-10-027-632-16037	Sequence 16037, A
C 59	23	48.9	683	17	US-10-027-632-16037	Sequence 16037, A
C 60	23	48.9	10329	15	US-10-027-632-16037	Sequence 2095, Ap
C 61	23	48.9	50000	17	US-10-364-505-6	Sequence 6, Appl1
C 62	23	48.9	50000	18	US-10-681-193-6	Sequence 11, Appl1
C 63	22.8	48.5	712	13	US-10-027-632-111416	Sequence 11, Appl1
C 64	22.8	48.5	712	17	US-10-027-632-111416	Sequence 11, Appl1
C 65	22.8	48.5	1223	13	US-10-027-632-213596	Sequence 213596, A
C 66	22.8	48.5	1223	17	US-10-027-632-213596	Sequence 213596, A
C 67	22.8	48.5	1506	17	US-10-335-977-4464	Sequence 4464, Ap
C 68	22.8	48.5	2751	17	US-10-335-977-4465	Sequence 4465, Ap
C 69	22.8	48.5	2784	17	US-10-335-977-4466	Sequence 4466, Ap
C 70	22.8	48.5	4161	18	US-10-437-963-87144	Sequence 87144, Ap
C 71	22.8	48.5	28397	18	US-10-322-696-43	Sequence 43, Appl1
C 72	22.8	48.5	31491	9	US-09-864-680-1	Sequence 1, Appl1
C 73	22.6	48.1	189	18	US-10-425-115-103413	Sequence 103413, A
C 74	22.6	48.1	431	17	US-10-424-559-139714	Sequence 139714, A
C 75	22.6	48.1	434	13	US-10-027-632-55719	Sequence 55719, A
C 76	22.6	48.1	434	13	US-10-027-632-309404	Sequence 309404, A
C 77	22.6	48.1	434	17	US-10-027-632-55719	Sequence 55719, A
C 78	22.6	48.1	434	17	US-10-027-632-309404	Sequence 309404, A
C 79	22.6	48.1	435	13	US-10-027-632-1985	Sequence 1985, Ap
C 80	22.6	48.1	435	17	US-10-027-632-1985	Sequence 1985, Ap
C 81	22.6	48.1	459	8	US-08-781-986A-1138	Sequence 1138, Ap
C 82	22.6	48.1	459	17	US-10-329-624-1138	Sequence 203374, A
C 83	22.6	48.1	611	13	US-10-027-632-203374	Sequence 203374, A
C 84	22.6	48.1	611	13	US-10-027-632-203375	Sequence 203375, A

C 85	22.6	48.1	611	17	US-10-027-632-203374	Sequence 203374,
C 86	22.6	48.1	611	17	US-10-027-632-203375	Sequence 203375,
C 87	22.6	48.1	701	18	US-10-653-047-6311	Sequence 6311, Ap
C 88	22.6	48.1	730	13	US-10-027-632-24249	Sequence 24249, A
C 89	22.6	48.1	730	17	US-10-027-632-24249	Sequence 24249, A
C 90	22.6	48.1	792	13	US-10-027-632-127002	Sequence 127002,
C 91	22.6	48.1	792	13	US-10-027-632-127003	Sequence 127003,
C 92	22.6	48.1	792	13	US-10-027-632-165053	Sequence 165053,
C 93	22.6	48.1	792	17	US-10-027-632-127002	Sequence 127002,
C 94	22.6	48.1	792	17	US-10-027-632-127003	Sequence 127003,
C 95	22.6	48.1	792	17	US-10-027-632-165053	Sequence 165053,
C 96	22.6	48.1	1209	9	US-09-939-980-55	Sequence 55, Appl
C 97	22.6	48.1	2319	18	US-10-437-963-87185	Sequence 87185, A
C 98	22.6	48.1	2526	17	US-10-425-114-31597	Sequence 31597, A
C 99	22.6	48.1	7325	17	US-10-052-482-196	Sequence 196, App
C 100	22.6	48.1	322101	16	US-10-060-902-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-10-349-143-1
; Sequence 1, Application US/10349143
; Publication No. US20040005584A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marla
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
; FILE REFERENCE: GENSET.020CPI
; CURRENT APPLICATION NUMBER: US/10/349,143
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US/09/422,978
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 09/298,850
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/109,732
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-23
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/082,614
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-21
; NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 1
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 24
; OTHER INFORMATION: 99-109-224 : polymorphic base G or C
US-10-349-143-1

Query Match 99.1%; Score 46.6; DB 17; Length 47;
Best Local Similarity 100.0%; Pred. No. 8.9e-06;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCACAGGCTTGATAGAAATAAATTTGATCACCATTTCATTTTCAATTTT 47
Db 1 CCACAGGCTTGATAGAAATAAATTTGATCACCATTTCATTTTCAATTTT 47

RESULT 2
US-09-790-289-1
; Sequence 1, Application US/09790289
; Publication No. US20030165826A1
; GENERAL INFORMATION:
; APPLICANT: Caroline Barry
; APPLICANT: Ilya Chumakov
; TITLE OF INVENTION: PG-3 and Biallelic Markers Thereof
; FILE REFERENCE: 68 US3 REG
; CURRENT APPLICATION NUMBER: US/09/790,289
; CURRENT FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 5

SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 240825
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: misc.feature
; NAME/KEY: 1..2000
; LOCATION: 1..2000
; OTHER INFORMATION: 5'regulatory region
; NAME/KEY: exon
; LOCATION: 2001..2079
; OTHER INFORMATION: exon A
; NAME/KEY: exon
; LOCATION: 4627..4718
; OTHER INFORMATION: exon B
; NAME/KEY: exon
; LOCATION: 10115..10233
; OTHER INFORMATION: exon C
; NAME/KEY: exon
; LOCATION: 26810..26897
; OTHER INFORMATION: exon D
; NAME/KEY: exon
; LOCATION: 31357..31471
; OTHER INFORMATION: exon E
; NAME/KEY: exon
; LOCATION: 34261..34404
; OTHER INFORMATION: exon F
; NAME/KEY: exon
; LOCATION: 37377..37466
; OTHER INFORMATION: exon S
; NAME/KEY: exon
; LOCATION: 39704..40858
; OTHER INFORMATION: exon T
; NAME/KEY: exon
; LOCATION: 50436..50545
; OTHER INFORMATION: exon G
; NAME/KEY: exon
; LOCATION: 72881..72918
; OTHER INFORMATION: exon H
; NAME/KEY: exon
; LOCATION: 75989..76151
; OTHER INFORMATION: exon I
; NAME/KEY: exon
; LOCATION: 95111..95188
; OTHER INFORMATION: exon J
; NAME/KEY: exon
; LOCATION: 216015..216252
; OTHER INFORMATION: exon K
; NAME/KEY: exon
; LOCATION: 237526..238825
; OTHER INFORMATION: exon L
; NAME/KEY: misc.feature
; LOCATION: 238826..240825
; OTHER INFORMATION: 3'regulatory region
; NAME/KEY: allele
; LOCATION: 1999
; OTHER INFORMATION: 5-390-177 : polymorphic base G or C
; NAME/KEY: allele
; LOCATION: 4601
; OTHER INFORMATION: 5-391-43 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 10228
; OTHER INFORMATION: 5-392-222 : polymorphic base G or T
; NAME/KEY: allele
; LOCATION: 10286
; OTHER INFORMATION: 5-392-280 : polymorphic base G or T
; NAME/KEY: allele
; LOCATION: 10370
; OTHER INFORMATION: 5-392-364 : insertion of G
; NAME/KEY: allele
; LOCATION: 39944
; OTHER INFORMATION: 4-58-318 : polymorphic base G or T
; NAME/KEY: allele

LOCATION: 39973
OTHER INFORMATION: 4-58-289 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 41385
OTHER INFORMATION: 4-54-199 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 41404
OTHER INFORMATION: 4-54-180 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 42232
OTHER INFORMATION: 4-51-312 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 67475
OTHER INFORMATION: 99-86-266 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 69521
OTHER INFORMATION: 4-88-107 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 72838
OTHER INFORMATION: 5-397-141 : polymorphic base G or T
NAME/KEY: allele
LOCATION: 76060
OTHER INFORMATION: 5-398-203 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 81253
OTHER INFORMATION: 99-12738-248 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 83921
OTHER INFORMATION: 99-109-358 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 91917
OTHER INFORMATION: 99-12749-175 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 95349
OTHER INFORMATION: 4-21-154 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 95511
OTHER INFORMATION: 4-21-317 : polymorphic base G or T
NAME/KEY: allele
LOCATION: 96190
OTHER INFORMATION: 4-23-326 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 97294
OTHER INFORMATION: 99-12753-34 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 98024
OTHER INFORMATION: 5-364-252 : polymorphic base G or T
NAME/KEY: allele
LOCATION: 98914
OTHER INFORMATION: 99-12755-280 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 98963
OTHER INFORMATION: 99-12755-329 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 103593
OTHER INFORMATION: 4-87-212 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 104398
OTHER INFORMATION: 99-12757-318 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 106373
OTHER INFORMATION: 99-12758-102 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 106407
OTHER INFORMATION: 99-12758-136 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 108315
OTHER INFORMATION: 4-105-98 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 108327
OTHER INFORMATION: 4-105-86 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 108472

OTHER INFORMATION: 4-45-49 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 109196
OTHER INFORMATION: 4-44-277 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 114604
OTHER INFORMATION: 4-86-60 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 115716
OTHER INFORMATION: 4-84-334 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 122083
OTHER INFORMATION: 99-78-321 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 123124
OTHER INFORMATION: 99-12767-36 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 123231
OTHER INFORMATION: 99-12767-143 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 123277
OTHER INFORMATION: 99-12767-189 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 123468
OTHER INFORMATION: 99-12767-380 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 126738
OTHER INFORMATION: 4-80-328 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 128210
OTHER INFORMATION: 4-36-384 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 128330
OTHER INFORMATION: 4-36-264 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 128333
OTHER INFORMATION: 4-36-261 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 128594
OTHER INFORMATION: 4-35-333 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 128687
OTHER INFORMATION: 4-35-240 : polymorphic base G or C
NAME/KEY: allele

Query Match 99.1%; Score 46.6; DB 10; Length 240825;
Best Local Similarity 97.9%; Pred. No. 8.4e-05;
Matches 46; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCACAGCCTTGATGAAATAASTTGTATCCCATTTTCAAAATTT 47
DB 83764 CCACAGCCTTGATGAAATAASTTGTATCCCATTTTCAAAATTT 83810

RESULT 3
US-10-468-582-1
Sequence 1, Application US/10468582
Publication No. US20040163137A1
GENERAL INFORMATION:
APPLICANT: Genset
TITLE OF INVENTION: PG-3 and biallelic markers thereof
FILE REFERENCE: 68. W02
CURRENT APPLICATION NUMBER: US/10/468,582
CURRENT FILING DATE: 2003-08-20
NUMBER OF SEQ ID NOS: 5
SOFTWARE: Patent.pm
SEQ ID NO 1
LENGTH: 240825
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..2000

OTHER INFORMATION: 5'regulatory region
FEATURE:
NAME/KEY: exon
LOCATION: 2001..2079
OTHER INFORMATION: exon A
FEATURE:
NAME/KEY: exon
LOCATION: 4627..4718
OTHER INFORMATION: exon B
FEATURE:
NAME/KEY: exon
LOCATION: 10115..10233
OTHER INFORMATION: exon C
FEATURE:
NAME/KEY: exon
LOCATION: 31357..31471
OTHER INFORMATION: exon E
FEATURE:
NAME/KEY: exon
LOCATION: 34261..34404
OTHER INFORMATION: exon F
FEATURE:
NAME/KEY: exon
LOCATION: 37377..37466
OTHER INFORMATION: exon S
FEATURE:
NAME/KEY: exon
LOCATION: 39704..40858
OTHER INFORMATION: exon T
FEATURE:
NAME/KEY: exon
LOCATION: 50436..50545
OTHER INFORMATION: exon G
FEATURE:
NAME/KEY: exon
LOCATION: 72881..72918
OTHER INFORMATION: exon H
FEATURE:
NAME/KEY: exon
LOCATION: 75989..76151
OTHER INFORMATION: exon I
FEATURE:
NAME/KEY: exon
LOCATION: 95111..95188
OTHER INFORMATION: exon J
FEATURE:
NAME/KEY: exon
LOCATION: 216015..216252
OTHER INFORMATION: exon K
FEATURE:
NAME/KEY: exon
LOCATION: 237526..238825
OTHER INFORMATION: exon L
FEATURE:
NAME/KEY: misc feature
LOCATION: 238826..240825
OTHER INFORMATION: 3'regulatory region
FEATURE:
NAME/KEY: allele
LOCATION: 1999
OTHER INFORMATION: 5-390-177 : polymorphic base G or C
FEATURE:
NAME/KEY: allele
LOCATION: 4601
OTHER INFORMATION: 5-391-43 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 10228
OTHER INFORMATION: 5-392-222 : polymorphic base G or T

FEATURE:
NAME/KEY: allele
LOCATION: 10286
OTHER INFORMATION: 5-392-280 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 10370
OTHER INFORMATION: 5-392-364 : insertion of G
FEATURE:
NAME/KEY: allele
LOCATION: 39944
OTHER INFORMATION: 4-58-318 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 39973
OTHER INFORMATION: 4-58-289 : polymorphic base G or C
FEATURE:
NAME/KEY: allele
LOCATION: 41385
OTHER INFORMATION: 4-54-199 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 41404
OTHER INFORMATION: 4-54-180 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 42232
OTHER INFORMATION: 4-51-312 : polymorphic base G or C
FEATURE:
NAME/KEY: allele
LOCATION: 67475
OTHER INFORMATION: 99-86-266 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 69521
OTHER INFORMATION: 4-88-107 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 72838
OTHER INFORMATION: 5-397-141 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 76060
OTHER INFORMATION: 5-398-203 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 81253
OTHER INFORMATION: 99-12738-248 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 83921
OTHER INFORMATION: 99-109-358 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 91917
OTHER INFORMATION: 99-12749-175 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 95349
OTHER INFORMATION: 4-21-154 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 95511
OTHER INFORMATION: 4-21-317 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 96190
OTHER INFORMATION: 4-23-326 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 97294
OTHER INFORMATION: 99-12753-34 : polymorphic base A or T
FEATURE:

NAME/KEY: allele
LOCATION: 98024
OTHER INFORMATION: 5-364-252 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 98914
OTHER INFORMATION: 99-12755-280 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 98963
OTHER INFORMATION: 99-12755-329 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 103593
OTHER INFORMATION: 4-87-212 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 104398
OTHER INFORMATION: 99-12757-318 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 106373
OTHER INFORMATION: 99-12758-102 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 106407
OTHER INFORMATION: 99-12758-136 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 108315
OTHER INFORMATION: 4-105-98 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 108327
OTHER INFORMATION: 4-105-86 : polymorphic base A or G
FEATURE:
NAME/KEY: allele

Query Match 99.1%; Score 46.6; DB 18; Length 240825;
Best Local Similarity 97.9%; Pred. No. 8.4e-05;
Matches 46; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCACGGCTGATGTAATTAATTTGATCCACATTTCAATTTT 47
DB 83764 CCACGGCTGATGTAATTAATTTGATCCACATTTCAATTTT 83810

RESULT 4
US-10-719-993-29983
Sequence 29983, Application US/10719993
Publication No. US20040265849A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
FILE REFERENCE: CL001496
CURRENT APPLICATION NUMBER: US/10/719,993
NUMBER OF SEQ ID NOS: 55342
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 29983
LENGTH: 201
TYPE: DNA
ORGANISM: Homo sapiens
US-10-719-993-29983

Query Match 52.8%; Score 24.8; DB 18; Length 201;
Best Local Similarity 76.3%; Pred. No. 1.9e+02;
Matches 29; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 10 TGATTGAATTAATTAATTTGATCCACATTTCAATTTT 47
DB 71 TTATAGAAATTAACCTTTAATTAACCTTTTAAATTT 108

RESULT 5
US-10-719-993-6843
Sequence 6843, Application US/10719993
Publication No. US20040265849A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
FILE REFERENCE: CL001496
CURRENT APPLICATION NUMBER: US/10/719,993
NUMBER OF SEQ ID NOS: 55342
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6843
LENGTH: 108111
TYPE: DNA
ORGANISM: Homo sapiens
US-10-719-993-6843

Query Match 52.8%; Score 24.8; DB 18; Length 108111;
Best Local Similarity 76.3%; Pred. No. 9.9e+02;
Matches 29; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 10 TGATTGAATTAATTAATTTGATCCACATTTCAATTTT 47
DB 11145 TTATAGAAATTAACCTTTAATTAACCTTTTAAATTT 11182

RESULT 6
US-09-796-692-4373/C
Sequence 4373, Application US/09796692
Publication No. US20020198362A1
GENERAL INFORMATION:
APPLICANT: Galger, Alexander
APPLICANT: Mannion, Jane

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
FILE REFERENCE: 2077.001200
CURRENT APPLICATION NUMBER: US/09/796,692
PRIOR FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: 60/223,378
PRIOR FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 9597
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4373
LENGTH: 267
TYPE: DNA
ORGANISM: Homo sapiens

US-09-796-692-4373

Query Match 51.9%; Score 24.4; DB 9; Length 267;
Best Local Similarity 70.5%; Pred. No. 2.8e+02;
Matches 31; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

Qy 4 CAGCGTTGATTAGAAATAAATTTGATCACCATTTCATTTT 47
Db 196 CATGCTTGCTTTTATCAACTGTATCAACCATTTTATTTT 153

RESULT 7

US-09-796-692-9383/c
; Sequence 9383, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9383
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-9383

Query Match 51.9%; Score 24.4; DB 9; Length 267;
Best Local Similarity 70.5%; Pred. No. 2.8e+02;
Matches 31; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

Qy 4 CAGCGTTGATTAGAAATAAATTTGATCACCATTTCATTTT 47
Db 196 CATGCTTGCTTTTATCAACTGTATCAACCATTTTATTTT 153

RESULT 8

US-10-040-862-4373/c
; Sequence 4373, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane

; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4373
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-4373

Query Match 51.9%; Score 24.4; DB 14; Length 267;
Best Local Similarity 70.5%; Pred. No. 2.8e+02;
Matches 31; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

Qy 4 CAGCGTTGATTAGAAATAAATTTGATCACCATTTCATTTT 47
Db 196 CATGCTTGCTTTTATCAACTGTATCAACCATTTTATTTT 153

RESULT 9

US-10-040-862-9383/c
; Sequence 9383, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28

```

; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 9383
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-040-862-9383

Query Match      51.9%; Score 24.4; DB 14; Length 267;
Best Local Similarity 70.5%; Pred. No. 2.8e+02;
Matches 31; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 4 CAGGCTGATTGAAATAASTTTGATCCACATTTCAATTTT 47
DB 196 CAGGCTGCTTTTATTCACAGTATACCATTTATTAATTTT 153

RESULT 10
; US-10-057-475B-4373/C
; Sequence 4373, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aljun
; APPLICANT: Ordenez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
```

```

; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 4373
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-057-475B-4373

Query Match      51.9%; Score 24.4; DB 17; Length 267;
Best Local Similarity 70.5%; Pred. No. 2.8e+02;
Matches 31; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 4 CAGGCTGATTGAAATAASTTTGATCCACATTTCAATTTT 47
DB 196 CAGGCTGCTTTTATTCACAGTATACCATTTATTAATTTT 153

RESULT 11
; US-10-057-475B-9383/C
; Sequence 9383, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aljun
; APPLICANT: Ordenez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 9383
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-057-475B-9383

Query Match      51.9%; Score 24.4; DB 17; Length 267;
Best Local Similarity 70.5%; Pred. No. 2.8e+02;
Matches 31; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 4 CAGGCTGATTGAAATAASTTTGATCCACATTTCAATTTT 47
DB 196 CAGGCTGCTTTTATTCACAGTATACCATTTATTAATTTT 153
```

DB 196 CAGCTTGCTTTTATCACTGTTATCAACATTATTAATTTT 153

RESULT 12

US-10-154-884B-4373/c

/ Sequence 4373, Application US/10154884B

/ Publication No. US20040005561A1

/ GENERAL INFORMATION:

/ APPLICANT: Gaiger, Alexander

/ APPLICANT: Algate, Paul A.

/ APPLICANT: Mannion, Jane

/ APPLICANT: Retter, Marc W.

/ APPLICANT: Corixa Corporation

/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy

/ FILE REFERENCE: 014058-013521US

/ CURRENT APPLICATION NUMBER: US/10/154,884B

/ PRIOR FILING DATE: 2002-05-23

/ PRIOR APPLICATION NUMBER: US 60/186,126

/ PRIOR FILING DATE: 2000-03-01

/ PRIOR APPLICATION NUMBER: US 60/190,479

/ PRIOR FILING DATE: 2000-03-17

/ PRIOR APPLICATION NUMBER: US 60/200,545

/ PRIOR FILING DATE: 2000-04-27

/ PRIOR APPLICATION NUMBER: US 60/200,303

/ PRIOR FILING DATE: 2000-04-28

/ PRIOR APPLICATION NUMBER: US 60/200,779

/ PRIOR FILING DATE: 2000-04-28

/ PRIOR APPLICATION NUMBER: US 60/200,999

/ PRIOR FILING DATE: 2000-05-01

/ PRIOR APPLICATION NUMBER: US 60/202,084

/ PRIOR FILING DATE: 2000-05-04

/ PRIOR APPLICATION NUMBER: US 60/206,201

/ PRIOR FILING DATE: 2000-05-22

/ PRIOR APPLICATION NUMBER: US 60/218,950

/ PRIOR FILING DATE: 2000-07-14

/ PRIOR APPLICATION NUMBER: US 60/222,903

/ PRIOR FILING DATE: 2000-08-03

/ Remaining Prior Application data removed - See File Wrapper or PALM.

/ NUMBER OF SEQ ID NOS: 11290

/ SOFTWARE: FastSeq for Windows Version 3.0

/ SEQ ID NO: 4373

/ LENGTH: 267

/ TYPE: DNA

/ ORGANISM: Homo sapiens

US-10-154-884B-4373

Query Match 51.9%; Score 24.4; DB 17; Length 267;

Best Local Similarity 70.5%; Pred. No. 2.8e+02;

Matches 31; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 4 CAGGCTGATTAGAAATAASTTTGATCCGATTTCAAATTTT 47

DB 196 CAGCTTGCTTTTATCACTGTTATCAACATTATTAATTTT 153

RESULT 13

US-10-154-884B-9383/c

/ Sequence 9383, Application US/10154884B

/ Publication No. US20040005561A1

/ GENERAL INFORMATION:

/ APPLICANT: Gaiger, Alexander

/ APPLICANT: Algate, Paul A.

/ APPLICANT: Mannion, Jane

/ APPLICANT: Retter, Marc W.

/ APPLICANT: Corixa Corporation

/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy

/ FILE REFERENCE: 014058-013521US

/ CURRENT APPLICATION NUMBER: US/10/154,884B

/ PRIOR FILING DATE: 2002-05-23

/ PRIOR APPLICATION NUMBER: US 60/186,126

/ PRIOR FILING DATE: 2000-03-01

/ PRIOR APPLICATION NUMBER: US 60/190,479

/ PRIOR FILING DATE: 2000-03-17

/ PRIOR APPLICATION NUMBER: US 60/200,545

/ PRIOR FILING DATE: 2000-04-27

/ PRIOR APPLICATION NUMBER: US 60/200,303

/ PRIOR FILING DATE: 2000-04-28

/ PRIOR APPLICATION NUMBER: US 60/200,779

/ PRIOR FILING DATE: 2000-04-28

/ PRIOR APPLICATION NUMBER: US 60/200,999

/ PRIOR FILING DATE: 2000-05-01

/ PRIOR APPLICATION NUMBER: US 60/202,084

/ PRIOR FILING DATE: 2000-05-04

/ PRIOR APPLICATION NUMBER: US 60/206,201

/ PRIOR FILING DATE: 2000-05-22

/ PRIOR APPLICATION NUMBER: US 60/218,950

/ PRIOR FILING DATE: 2000-07-14

/ PRIOR APPLICATION NUMBER: US 60/222,903

/ PRIOR FILING DATE: 2000-08-03

/ Remaining Prior Application data removed - See File Wrapper or PALM.

/ NUMBER OF SEQ ID NOS: 11290

/ SOFTWARE: FastSeq for Windows Version 3.0

/ SEQ ID NO: 9383

/ LENGTH: 267

/ TYPE: DNA

/ ORGANISM: Homo sapiens

US-10-154-884B-9383

Query Match 51.9%; Score 24.4; DB 17; Length 267;

Best Local Similarity 70.5%; Pred. No. 2.8e+02;

Matches 31; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 4 CAGGCTGATTAGAAATAASTTTGATCCGATTTCAAATTTT 47

DB 196 CAGCTTGCTTTTATCACTGTTATCAACATTATTAATTTT 153

RESULT 14

US-10-764-324-4373/c

/ Sequence 4373, Application US/10764324

/ Publication No. US20040175739A1

/ GENERAL INFORMATION:

/ APPLICANT: Gaiger, Alexander

/ APPLICANT: Algate, Paul A.

/ APPLICANT: Mannion, Jane

/ APPLICANT: Retter, Marc

/ APPLICANT: Corixa Corporation

/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy

/ FILE REFERENCE: 014058-013520US

/ CURRENT APPLICATION NUMBER: US/10/764,324

/ PRIOR FILING DATE: 2004-01-23

/ PRIOR APPLICATION NUMBER: US/10/040,862

/ PRIOR FILING DATE: 2001-11-06

/ PRIOR APPLICATION NUMBER: US 60/186,126

/ PRIOR FILING DATE: 2000-03-01

/ PRIOR APPLICATION NUMBER: US 60/190,479

/ PRIOR FILING DATE: 2000-03-17

/ PRIOR APPLICATION NUMBER: US 60/200,545

/ PRIOR FILING DATE: 2000-04-27

/ PRIOR APPLICATION NUMBER: US 60/200,303

/ PRIOR FILING DATE: 2000-04-28

/ PRIOR APPLICATION NUMBER: US 60/200,779

/ PRIOR FILING DATE: 2000-04-28

/ PRIOR APPLICATION NUMBER: US 60/200,999

/ PRIOR FILING DATE: 2000-05-01

/ PRIOR APPLICATION NUMBER: US 60/202,084

/ PRIOR FILING DATE: 2000-05-04

/ PRIOR APPLICATION NUMBER: US 60/206,201

/ PRIOR FILING DATE: 2000-05-22

/ PRIOR APPLICATION NUMBER: US 60/218,950

/ PRIOR FILING DATE: 2000-07-14

/ Remaining Prior Application data removed - See File Wrapper or PALM.

/ NUMBER OF SEQ ID NOS: 10467

SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4373
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-764-324-4373

Query Match 51.9%; Score 24.4; DB 18; Length 267;
Best Local Similarity 70.5%; Pred. No. 2.8e+02;
Matches 31; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

Qy 4 CAGGCTTGATTGAAATAAATTTGATCACCATTTCCTCAATTTT 47
Db 196 CATGCTTGCTTTTTCATCACTGTATCACCATTATTAATTTT 153

RESULT 15
US-10-764-324-9383/c
; Sequence 9383, Application US/10764324
; Publication No. US20040175739A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/764,324
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9383
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-764-324-9383

Query Match 51.9%; Score 24.4; DB 18; Length 267;
Best Local Similarity 70.5%; Pred. No. 2.8e+02;
Matches 31; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

Qy 4 CAGGCTTGATTGAAATAAATTTGATCACCATTTCCTCAATTTT 47
Db 196 CATGCTTGCTTTTTCATCACTGTATCACCATTATTAATTTT 153

RESULT 16
US-10-027-632-240198/c
; Sequence 240198, Application US/10027632
; Publication No. US20020198371A1

GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 240198
; LENGTH: 589
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-240198

Query Match 51.9%; Score 24.4; DB 13; Length 589;
Best Local Similarity 70.5%; Pred. No. 3.4e+02;
Matches 31; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

Qy 4 CAGGCTTGATTGAAATAAATTTGATCACCATTTCCTCAATTTT 47
Db 553 CATGCTTGCTTTTTCATCACTGTATCACCATTATTAATTTT 510

RESULT 17
US-10-027-632-240198/c
; Sequence 240198, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 240198
; LENGTH: 589
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-240198

Query Match 51.9%; Score 24.4; DB 13; Length 589;
Best Local Similarity 70.5%; Pred. No. 3.4e+02;

PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 240200
LENGTH: 589
TYPE: DNA
ORGANISM: Human
US-10-027-632-240200

Query Match 51.9%; Score 24.4; DB 17; Length 589;
Best Local Similarity 70.5%; Pred. No. 3.4e+02;
Matches 31; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 4 CAGGCTTGATTGAATAAATTAATTTGATCACCATTTCATTTCAATTTT 47
DB 553 CAGGCTTGCTTTTTCATCACTGTTATCCACATTTTATTAATTTT 510

RESULT 22
US-10-437-963-51548
Sequence 51548, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 51548
LENGTH: 2199
TYPE: DNA
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_5392C.1
US-10-437-963-51548

Query Match 51.5%; Score 24.2; DB 18; Length 2199;
Best Local Similarity 71.1%; Pred. No. 5.6e+02;
Matches 32; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 CCACGGCTTGATTGAATAAATTAATTTGATCACCATTTCATTTCAATTT 45
DB 82 CCACGGCTGATTGACACAAATTTTCATGATTATTTCCACACT 126

RESULT 23
US-10-311-455-924/C
Sequence 924, Application US/10311455
Publication No. US20030143606A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
APPLICANT: BERLIN, Kurt

TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determi
FILE OF INVENTION: cytosine methylation
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311,455
CURRENT FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: PCT/EP01/07537
PRIOR FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 924
LENGTH: 5173
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
FEATURE:
NAME/KEY: unsure
LOCATION: 2436, 2554, 2736, 3004, 4504, 4580
OTHER INFORMATION: n is a or g or c or t
US-10-311-455-924

Query Match 51.5%; Score 24.2; DB 15; Length 5173;
Best Local Similarity 71.1%; Pred. No. 7e+02;
Matches 32; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 CCACAGGCTTGATTGAATAAATTAATTTGATCACCATTTCATTTCAATTT 45
DB 3820 CCATATCTTAATTAAATAATTAATTTCACTAATTTCTTAATTT 3776

RESULT 24
US-10-767-701-2646
Sequence 2646, Application US/10767701
Publication No. US20040172684A1
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Li, Ping
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 2646
LENGTH: 728
TYPE: DNA
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS44741_1
US-10-767-701-2646

Query Match 50.6%; Score 23.8; DB 18; Length 728;
Best Local Similarity 75.7%; Pred. No. 5.7e+02;
Matches 28; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 11 GATTGAATAAATTAATTTGATCACCATTTCATTTCAATTTT 47
DB 686 GATTAAACACAGGTTATTTAAACATTTTCATTTTCAATTTT 722

RESULT 25
US-10-425-114-15544
Sequence 15544, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; NUMBER OF SEQ ID NOS: 2003-04-28
 ; SEQ ID NO 15544
 ; LENGTH: 1102
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: LIB3060-021-G2_FLI
 US-10-425-114-15544

Query Match 50.6%; Score 23.8; DB 17; Length 1102;
 Best Local Similarity 75.7%; Pred. No. 6.3e+02;
 Matches 28; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Oy 11 GATTGAATATAAATTGATCACCATTTCATAATTTT 47
 |||||:|||||||
 Db 625 GAGTGATATATAAGTGTCTCATCCATATTACCGTTT 661

Search completed: June 2, 2005, 10:31:30
 Job time : 525 secs